



**Tiago Henrique da Costa Veloso**

Licenciado em Bioquímica

## **Fosforilação do efetor Tir e a sua interação com proteínas alvo do hospedeiro**

Dissertação para obtenção do Grau de Mestre em  
Bioquímica

Orientador: Doutor Tiago Neto Cordeiro, Investigador auxiliar, ITQB-NOVA

Júri:

Presidente: Professora Doutora Sofia Pauleta

Arguente: Professor Doutor Eurico Cabrita

Vogal: Doutor Tiago Cordeiro

**Dezembro, 2020**



FACULDADE DE  
CIÊNCIAS E TECNOLOGIA  
UNIVERSIDADE NOVA DE LISBOA



**Tiago Henrique da Costa Veloso**

Licenciado em Bioquímica

**Fosforilação do efetor Tir e a sua interação com  
proteínas alvo do hospedeiro**

Dissertação para obtenção do Grau de Mestre em  
Bioquímica

Orientador: Doutor Tiago Neto Cordeiro, Investigador auxiliar, ITQB-NOVA

Júri:

Presidente: Professora Doutora Sofia Pauleta

Arguente: Professor Doutor Eurico Cabrita

Vogal: Doutor Tiago Cordeiro

**Dezembro, 2020**



FACULDADE DE  
CIÊNCIAS E TECNOLOGIA  
UNIVERSIDADE NOVA DE LISBOA



### **Fosforilação do efetor Tir e a sua interação com proteínas alvo do hospedeiro**

Copyright © Tiago Henrique da Costa Veloso, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.



# Agradecimentos

Em primeiro lugar, quero agradecer ao meu orientador Doutor Tiago Cordeiro, alguém que admiro pelo seu progresso científico e a quem quero agradecer a oportunidade de desenvolver a dissertação no seu grupo de investigação. Por me ter transmitido conhecimento e por ter incutido em mim espírito crítico e científico. Estou muito grato pela experiência enriquecedora que me proporcionou.

A todos os elementos do grupo *Dynamic Structural Biology*, deixo o meu profundo agradecimento. Ao Guillem quero agradecer pela ajuda nos momentos de dúvida e por nunca me deixar desmotivar mesmo quando as coisas corriam menos bem. Em especial quero agradecer à Marta, pela paciência e preocupação em fazer com que eu entendesse tudo o que estava a fazer e por exigir todo o meu empenho e dedicação. Se neste momento me considero alguém com capacidades para trabalhar num laboratório de investigação, devo-to a ti! Um obrigado enorme!

Ao ITQB e à FCT-UNL, por proporcionar as condições necessárias ao desenvolvimento da presente dissertação. Agradecer também aos técnicos da instituição que, de alguma forma, também foram intervenientes no trabalho desenvolvido, em especial à Doutora Cristina Timóteo por toda a disponibilidade e pelos conhecimentos transmitidos.

Ao grupo de amigos que se formou este ano Anaísa, Carolina, Diogo, Jéssica, Marta e Rodrigo, obrigado pelos almoços/lanches, pelas conversas sempre cheias de conteúdo (ou não) que permitiam o distanciamento suficiente do trabalho de modo a relaxar e voltar ao mesmo com outra energia. Um especial agradecimento à Anaísa, por ver sempre o lado positivo das situações e por toda a disponibilidade demonstrada, mesmo nos dias em que andava literalmente a correr.

Aos amigos de uma vida Filipe, Galo, Gonçalo, João, Lima, Ricardo, Tati, Zé e aos que não são de uma vida, mas que são igualmente importantes André, Carlos, Bia, Erikson, Matias, Mesquita, Thomas, Raquel, Rodrigo, Rui e Vidinha obrigado por todos os momentos de descontração e distração, pelo apoio, pela ramboia... de uma maneira mais simples: obrigado por tornarem tudo mais fácil. Vocês sabem quem são e o que significam!

À Tatiana, obrigado por tudo. Obrigado pelo sorriso que dá sentido a tudo.

Um grande obrigado à minha família, em especial aos meus pais Américo e Isabel pela paciência em ouvir todos os meus problemas quando as coisas corriam menos bem. Por me motivarem e fazerem de mim a pessoa que sou hoje. Um obrigado à pequena Leonor, que sem saber, tornou este ano mais especial.





## Resumo

Tir é uma proteína efetora intrinsecamente desordenada (IDP) de *Escherichia coli* Enteropatogénica (EPEC) e Enterohemorrágica (EHEC) que é translocada para o interior do hospedeiro. Na região C-terminal (C-Tir), este efector mimetiza quatro motivos, conhecidos como ITIMs, que contêm tirosinas (Y454, Y474, Y483, Y511) que são fosforiladas por cinases do hospedeiro. Os padrões de fosforilação resultantes são reconhecidos por uma série de proteínas do hospedeiro que desencadeiam mudanças na sinalização e nos processos celulares do mesmo. Destas proteínas, destaca-se a fosfatase SHP-1, essencial em vias de sinalização de resposta imunitária.

Nesta dissertação propusemos clarificar os aspetos moleculares da interação entre C-Tir e o domínio C-SH2 de SHP-1 e perceber como molecularmente a fosforilação molda essa interação.

Através da espectroscopia de RMN, foi realizada a reatribuição das ressonâncias da versão fosforilada de C-Tir e foram analisadas as propensões de estruturas secundárias. Este estudo permitiu a identificação de regiões com tendência para hélices- $\alpha$ , localizadas em zonas conhecidas por interagir com proteínas do hospedeiro.

A fosforilação é descrita como um requisito para a interação entre C-Tir e SHP-1. Contudo, os ensaios de interação realizados entre as duas proteínas pré-fosforilação sugerem que a mesma pode ocorrer intrinsecamente, através da região ITIM, que compreende Y511, identificada como uma região com preponderância para estrutura hélice- $\alpha$ .

Foi monitorizada a cinética e a extensão da fosforilação das quatro tirosinas, demonstrando que todas são fosforiladas simultaneamente, com Y454 a ser a mais rápida e Y483 a ser a mais lenta. Além disso, foi verificado que C-Tir fosforilada interage com o domínio C-SH2, através das quatro regiões ITIM. A região ITIM que contém Y454, cuja fosforilação é mais rápida, é a que apresenta menos afinidade para C-SH2, comparativamente com as restantes que apresentam afinidades semelhantes entre elas.

Este trabalho abre caminho para novos estudos sobre os mecanismos de virulência com implicações na saúde e na doença.

**Palavras-chave:** Tir; SHP-1; modificações pós-traducionais; interação proteína-proteína; proteínas intrinsecamente desordenadas; espectroscopia de RMN.



# Abstract

Tir is an intrinsically disordered protein (IDP) from Enteropathogenic (EPEC) and Enterohemorrhagic (EHEC) *Escherichia coli* that is translocated into the host. In its C-terminal region (C-Tir), this effector mimics four motifs, commonly known as ITIMs, containing tyrosine residues (Y454, Y474, Y483, Y511) that are phosphorylated by host kinases. The resulting phosphorylation patterns are recognized by various host proteins that trigger changes in signaling and host cellular processes. Among these proteins is SHP-1 phosphatase, an essential protein in signaling pathways of the immune response.

In this thesis, we proposed clarifying the molecular aspects of the interaction between C-Tir and the C-SH2 domain of SHP-1 and understanding how phosphorylation shapes this interaction.

Using NMR spectroscopy, we reassigned the backbone NMR resonances of the phosphorylated version of C-Tir and analyzed its secondary structural propensities. This study allowed us to identify regions with a tendency for helical structures located at sites known to interact with host proteins.

Phosphorylation seems to be essential for the interaction between Tir and SHP-1. However, binding assays between the two proteins pre-phosphorylation suggests that it can occur intrinsically through the ITIM region comprising Y511, identified to adopt a partial-folded  $\alpha$ -helix

We also monitored the kinetics and the extension of the phosphorylation reaction of the four tyrosines in the ITIM regions. In doing so, we found that the four positions are phosphorylated simultaneously, with Y454 being the fastest and Y483 being the slowest. Moreover, we showed that phosphorylated C-Tir interacts with the C-SH2 via its four ITIMs. The ITIM region containing Y454, whose phosphorylation is faster, is the one with a lower affinity for the C-SH2, than the others with similar binding strengths.

This work paves the way for further studies to understand virulence mechanisms with implications for health and disease.

**Keywords:** Tir; SHP-1; post-translational modifications; protein-protein interaction; intrinsically disordered proteins NMR spectroscopy.



# Prefácio

O que determina se a bactéria é eliminada ou se sobrevive e se multiplica infectando o hospedeiro é o equilíbrio entre os seus mecanismos patogênicos bacterianos e a resposta imunitária do hospedeiro. A presente dissertação aborda a interação entre duas proteínas, uma bacteriana e outra do hospedeiro, manifestamente importante para o inclinar da balança deste equilíbrio a favor da infecção, uma vez que esta interação “patógeno-hospedeiro” a nível molecular, permite a certas estirpes de *Escherichia Coli* patogênicas como EPEC e EHEC suprimirem a resposta imune do hospedeiro.

O documento está organizado em 7 grandes capítulos. O primeiro é a Introdução, que por sua vez está dividida em duas secções. Na primeira destas secções será abordada de uma forma geral a patogenicidade bacteriana de EPEC e EHEC e a estratégia por estas desenvolvida para controlar o hospedeiro e levar a cabo a infecção. Será dado o destaque para as duas proteínas estudadas na presente dissertação: o efector de virulência Tir libertado no interior do hospedeiro pelos patógenos EPEC e EHEC; e a fosfatase SHP-1, envolvida em diversos caminhos de sinalização da resposta imunitária. A espectroscopia de RMN foi a técnica utilizada ao longo do trabalho. Sendo o efector Tir uma proteína intrinsecamente desordenada (IDP) e sabendo que a espectroscopia RMN é por excelência o método mais adequado para estudar a dinâmica estrutural deste tipo de proteínas, na segunda secção deste capítulo serão abordados aspetos relevantes sobre esta classe de proteínas e será discutida a sua caracterização através da técnica referida. Será de igual forma discutida uma das modificações pós-traducionais mais comuns em IDPs, a fosforilação, e a forma como esta técnica é utilizada na sua monitorização. Esta secção trata ainda de interações entre macromoléculas, nomeadamente interações proteína-proteína, onde serão discutidos e explicados os diferentes regimes de permuta e a forma como se retiram conclusões acerca de cada um deles.

Depois de um pequeno capítulo onde os objetivos da presente dissertação se encontram definidos, segue-se um capítulo de Materiais e Métodos com os procedimentos experimentais realizados e com os detalhes do tratamento dos dados obtidos.

O capítulo seguinte aborda os Resultados e Discussão, onde como o próprio nome indica, serão apresentados e discutidos os resultados obtidos ao longo da presente dissertação. Este capítulo tal como o anterior, encontra-se dividido em duas secções. Uma primeira na qual será abordada a expressão e purificação das proteínas recombinantes e uma segunda referente à caracterização das mesmas utilizando a espectroscopia de RMN.

A esse capítulo, segue-se um referente às Conclusões e perspetivas futuras, na qual serão descritos os principais resultados e onde estão sugeridas algumas abordagens a realizar no futuro, com vista a alcançar outros objetivos inicialmente pensados e propostos. Os capítulos finais contêm a Bibliografia e os Anexos. Os últimos contêm informação complementar ao trabalho realizado e dados sobre os quais alguns resultados obtidos foram derivados.



# Índice de conteúdos

|   |       |
|---|-------|
| Agradecimentos.....   | V     |
| Resumo .....  | VII   |
| Abstract .....  | IX    |
| Prefácio .....  | XI    |
| Índice de figuras .....   | XVII  |
| Índice de tabelas .....   | XXI   |
| Lista de Abreviaturas.....  | XXIII |
| 1. Introdução .....   | 1     |
| 1.1. Virulência bacteriana e controlo do hospedeiro .....   | 3     |
| 1.1.1. Patogenicidade bacteriana de <i>Escherichia coli</i> enteropatogénica e enterohemorrágica (EPEC/EHEC).....   | 3     |
| 1.1.2. Recetor de intimina translocado (Tir) – o mestre da virulência.....  | 6     |
| 1.1.2.1. Interação entre Tir e proteínas com domínios SH2 – estratégia de infeção e inibição de resposta imune..... | 8     |
| 1.1.3. Fosfatase com domínio de homologia Src-2-1 (SHP-1) .....   | 11    |
| 1.2. Estrutura, dinâmica e interação de Proteínas Intrinsecamente Desordenadas (IDPs) .....                         | 13    |
| 1.2.1. IDPs .....   | 14    |
| 1.2.2. Ressonância Magnética Nuclear em IDPs.....   | 15    |
| 1.2.2.1. Reação de fosforilação <i>in vitro</i> monitorizada através da espectroscopia de RMN .....                 | 19    |
| 1.2.2.2. Fosforilação <i>in vitro</i> monitorizada em tempo real contínuo e descontínuo.....                        | 20    |
| 1.2.2.3. Caracterização de interações proteína-proteína através de espectroscopia de RMN .....                      | 21    |
| 2. Objetivos .....  | 27    |
| 3. Materiais e Métodos.....   | 31    |
| 3.1. Produção de proteína recombinante.....   | 33    |
| 3.1.1. Proteína C-Tir .....   | 33    |
| 3.1.1.1 Expressão da proteína isotopicamente marcada .....  | 33    |
| 3.1.1.2 Purificação da proteína isotopicamente marcada .....  | 34    |
| 3.1.2. Proteína SHP-1 .....   | 35    |

|          |  |    |
|----------|--|----|
| 3.1.2.1. | Clonagem dos genes dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1 .....                         | 35 |
| 3.1.2.2. | Testes de expressão dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1 .....                        | 38 |
| 3.1.2.3. | Expressão dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1 .....                                  | 40 |
| 3.1.2.4. | Purificação do domínio C-SH2 de SHP-1 .....  | 41 |
| 3.1.2.5. | Tentativas de purificação dos domínios N-SH2 e NC-SH2 de SHP-1 .....                         | 42 |
| 3.2.     | Caracterização de C-Tir através da espectroscopia de RMN .....                               | 42 |
| 3.2.1.   | Atribuição das ressonâncias de C-Tir fosforilada .....                                       | 43 |
| 3.2.1.1. | Preparação da amostra e aquisição dos dados .....  | 43 |
| 3.2.1.2. | Processamento dos dados .....  | 44 |
| 3.2.2.   | Reação de fosforilação de C-Tir <i>in vitro</i> monitorizada em tempo real descontínuo ..... | 44 |
| 3.2.2.1. | Preparação da amostra e aquisição dos dados .....  | 44 |
| 3.2.2.2. | Processamento dos dados .....  | 45 |
| 3.2.3.   | Interação entre C-Tir e o domínio C-SH2 de SHP-1 .....                                       | 46 |
| 3.2.3.1. | Preparação da amostra e aquisição dos dados .....  | 46 |
| 3.2.3.2. | Processamento dos dados .....  | 46 |
| 3.2.4.   | Interação entre C-Tir fosforilada e o domínio C-SH2 de SHP-1 .....                           | 47 |
| 3.2.4.1. | Preparação da amostra e aquisição dos dados .....  | 48 |
| 3.2.4.2. | Processamento dos dados .....  | 48 |
| 4.       | Resultados e Discussão .....   | 49 |
| 4.1.     | Produção de proteína recombinante .....  | 51 |
| 4.1.1.   | Proteína C-Tir .....   | 51 |
| 4.1.1.1. | Expressão e purificação da proteína isotopicamente marcada .....                             | 51 |
| 4.1.2.   | Proteína SHP-1 .....   | 52 |
| 4.1.2.1. | Clonagem dos genes dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1 .....                         | 52 |
| 4.1.2.2. | Testes de expressão dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1 .....                        | 56 |
| 4.1.2.3. | Expressão e purificação do domínio C-SH2 de SHP-1 .....                                      | 58 |
| 4.1.2.4. | Expressão e tentativas de purificação dos domínios N-SH2 e NC-SH2 de SHP-1 .....             | 60 |
| 4.2.     | Caracterização de C-Tir através da espectroscopia de RMN .....                               | 62 |
| 4.2.1.   | Atribuição das ressonâncias de C-Tir fosforilada .....                                       | 63 |
| 4.2.2.   | Propensão de elementos de estrutura secundária de C-Tir fosforilada .....                    | 65 |



|       |   |    |
|-------|---|----|
| 4.2.3 | Reação de fosforilação de C-Tir <i>in vitro</i> monitorizada em tempo real descontínuo..... | 67 |
| 4.2.4 | Interação entre C-Tir e o domínio C-SH2 de SHP-1 .....                                      | 71 |
| 4.2.5 | Interação entre C-Tir fosforilada e o domínio C-SH2 de SHP-1 .....                          | 74 |
| 5.    | Conclusões e perspectivas futuras.....  | 77 |
| 6.    | Bibliografia.....   | 83 |
| 7.    | Anexos .....  | 95 |



## Índice de figuras

|   |    |
|---|----|
| <b>Figura 1.1-</b> Esquema representativo do modelo de patogenicidade de EPEC/EHEC mediada por Tir. ....  | 6  |
| <b>Figura 1.2-</b> Propensão da desordem estrutural de Tir dos agentes patogénicos EPEC e EHEC. ....  | 7  |
| <b>Figura 1.3-</b> Mimetização de regiões ITIM e controlo do hospedeiro por Tir EPEC.....   | 10 |
| <b>Figura 1.4-</b> Estrutura geral de SHP-1.....  | 12 |
| <b>Figura 1.5-</b> Atribuição das ressonâncias da cadeia principal de uma proteína com recurso a experiências 3D de RMN. ....   | 17 |
| <b>Figura 1.6-</b> Fosforilação em resíduos de Serina, Treonina e Tirosina.....   | 19 |
| <b>Figura 1.7-</b> Esquema geral dos procedimentos de RMN para monitorizar reação de fosforilação de proteína in vitro em tempo real.....   | 21 |
| <b>Figura 1.8-</b> Diagrama esquemático dos espetros 2D <sup>1</sup> H- <sup>15</sup> N HSQC de uma titulação de um complexo, da esquerda para a direita, da proteína marcada A com concentração crescente de proteína B (ligando não marcado)..... | 24 |
| <b>Figura 4.1-</b> SDS-PAGE (12%) da purificação de C-Tir.. ....  | 52 |
| <b>Figura 4.2-</b> Esquema ilustrativo dos princípios da clonagem por reação LIC.....   | 54 |
| <b>Figura 4.3-</b> Géis de agarose 1% (m/v) da amplificação dos genes de interesse dos domínios de SHP-1.....   | 55 |
| <b>Figura 4.4-</b> Géis de agarose 1% (m/v) resultantes do screening das colónias recombinantes obtidas depois da reação de clonagem dos genes dos domínios de SHP-1.. ....   | 56 |
| <b>Figura 4.5-</b> SDS-PAGE (12%) dos testes de expressão dos domínios N-SH2 nos vetores pHTP8 (A) e pHTP10 (B), C-SH2 no vetor pHTP8 (C) e NC-SH2 nos vetores pHTP8 (D) e pHTP10 (E). ....   | 57 |

|   |     |
|---|-----|
| <b>Figura 4.6-</b> SDS-PAGE (12%) com as frações provenientes do primeiro passo de purificação do domínio C SH2 de SHP-1 por cromatografia de afinidade..                                     | 58  |
| <b>Figura 4.7-</b> SDS-PAGE (12%) referentes à clivagem enzimática do domínio C-SH2 com a protease HRV-3C e com as frações provenientes do segundo passo cromatográfico, através da IMAC..... | 59  |
| <b>Figura 4.8-</b> Cromatograma e SDS-PAGE (12%) obtidos no último passo de purificação do domínio C SH2, realizado por cromatografia de exclusão molecular..                                 | 60  |
| <b>Figura 4.9-</b> SDS-PAGE (12%) da tentativa de purificação da construção Trx-His6-NC-SH2....   | 61  |
| <b>Figura 4.10-</b> Atribuição das ressonâncias de C-Tir fosforilada. ....  | 64  |
| <b>Figura 4.11-</b> Desvios químicos secundários ( $\Delta\delta$ ) calculados para C-Tir no estado fosforilado.  | 66  |
| <b>Figura 4.12-</b> Reação de fosforilação de C-Tir in vitro monitorizado por RMN em tempo real descontínuo. ....   | 69  |
| <b>Figura 4.13-</b> Estudo da interação entre C-Tir e o domínio C-SH2 de SHP-1. ....  | 72  |
| <b>Figura 4.14-</b> Estudo da interação entre C-Tir fosforilada e o domínio C-SH2 de SHP-1....  | 75  |
| <b>Figura 5.1-</b> Representação esquemática das conclusões retiradas a partir do trabalho desenvolvido na presente dissertação.....  | 82  |
| <b>Figura 7.1-</b> Vetores de expressão nos quais foram clonados os genes das proteínas C-Tir, N-SH2, C-SH2 e NC-SH2 .....  | 97  |
| <b>Figura 7.2-</b> SDS-PAGE (12%) da tentativa de purificação da construção Trx-His6-N-SH2. ...   | 102 |
| <b>Figura 7.3-</b> SDS-PAGE (12%) da tentativa de purificação da construção Trx-His6-NC-SH2 com as frações provenientes da IMAC realizada utilizando a coluna HisTrap™. ....                  | 103 |
| <b>Figura 7.4-</b> SDS-PAGE (12%) da tentativa de purificação das construções NusA-His6-N-SH2 e NusA His6-NC-SH2.....   | 103 |
| <b>Figura 7.5-</b> SDS-PAGE (12%) da tentativa de purificação da construção Trx-His6-L8-NC-SH2..  | 104 |

**Figura 7.6-** Desvios químicos secundários ( $\Delta\delta C\alpha$  e  $\Delta\delta C\beta$ ) calculados para C-Tir no estado fosforilado..... 104

**Figura 7.7-** Curvas de titulação dos resíduos vizinhos e compreendidos nas quatro regiões ITIM presentes na sequência de C-Tir. .... 105



## Índice de tabelas

|   |    |
|---|----|
| <b>Tabela 3.1-</b> Construção do gene C-Tir EPEC no vetor pHTP8. ....   | 33 |
| <b>Tabela 3.2-</b> Construções dos genes C-SH2, NC-SH2 e SHP-1 WT nos vetores pGEX6P1 e pGEX-2T.....  | 35 |
| <b>Tabela 3.3-</b> Construções realizadas com os genes de interesse: N-SH2, C-SH2 e NC-SH2 de SHP-1.....  | 38 |
| <b>Tabela 3.4-</b> Condições testadas para a sobre-expressão das diferentes proteínas de interesse: domínio N-SH2, C-SH2 e NC-SH2 de SHP-1.....   | 39 |
| <b>Tabela 3.5-</b> Condições testadas nas diversas tentativas de purificação dos domínios N-SH2 e NC-SH2 de SHP-1. ....   | 42 |
| <b>Tabela 3.6-</b> Parâmetros das experiências de RMN.....  | 43 |
| <b>Tabela 4.1-</b> Condições otimizadas para a expressão dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1 na fração solúvel.....   | 57 |
| <b>Tabela 7.1-</b> Informação das proteínas C-Tir, N-SH2, C-SH2 e NC-SH2 abordadas nesta dissertação. ....  | 97 |
| <b>Tabela 7.2-</b> Programas de amplificação dos genes N-SH2, C-SH2, NC-SH2 de SHP-1 para clonagem das diferentes construções nos vetores pHTP8 pHTP10 e pCoofy18, através de PCR. ....         | 98 |
| <b>Tabela 7.3-</b> Componentes da amplificação dos genes N-SH2, C-SH2, NC-SH2 de SHP-1 para clonagem das diferentes construções nos vetores pHTP8, pHTP10 e pCoofy18, através de PCR. ....      | 98 |
| <b>Tabela 7.4-</b> Componentes e programa da reação de clonagem dos genes N-SH2, C-SH2, NC-SH2 de SHP-1 para clonagem das diferentes construções nos vetores pHTP8 pHTP10, através de PCR. .... | 99 |
| <b>Tabela 7.5-</b> Componentes e programa da reação de clonagem dos genes N-SH2 e NC-SH2 de SHP-1 para clonagem no vetor pCoofy18, através de PCR. ....   | 99 |
| <b>Tabela 7.6-</b> Programas do screening das colônias recombinantes dos genes N-SH2, C-SH2, NC-SH2 de SHP-1 após clonagem das diferentes construções nos vetores pHTP8 pHTP10.....             | 99 |

|   |     |
|---|-----|
| <b>Tabela 7.7-</b> Componentes do screening das colónias recombinantes dos genes alvo N-SH2, C-SH2, NC-SH2 de SHP-1 após clonagem das diferentes construções nos vetores pHTP8 pHTP10. .... | 100 |
| <b>Tabela 7.8-</b> Primers utilizados na amplificação dos genes alvo N-SH2, C-SH2, NC-SH2.....  | 100 |
| <b>Tabela 7.9-</b> Composição do gel SDS-PAGE (12%).....  | 101 |
| <b>Tabela 7.10-</b> Composição do tampão de corrida Tris-Glicina para SDS-PAGE pH=8,3 (1L)..  | 101 |
| <b>Tabela 7.11-</b> Composição do gel de agarose 1% (m/v). ....   | 101 |
| <b>Tabela 7.12-</b> Meio de cultura mínimo M9 para expressão de proteínas recombinantes isotopicamente marcadas. ....   | 102 |



# Lista de Abreviaturas

**Abs-** Absorvância

**APS-** Persulfato de amônia

**ATP-** Adesina trifosfato

**CD-** Dicroísmo Circular

**DNA** – Ácido desoxirribonucleico (do inglês, *deoxyribonucleic acid*)

**DTT** – Ditioneitol

**EDTA** – Ácido etilenodiamino tetra-acético

**EPEC-** *Escherichia coli* Enteropatogénica

**EHEC-** *Escherichia coli* Enterohemorrágica

**HEPES** – (4-(2-hidroxietil) - ácido 1-piperazinoetanosulfónico)

**His<sub>6</sub>** – Seis resíduos de Histidina

**IDP-** Proteína intrinsecamente desordenada (do inglês, inglês *Intrinsically Disordered Protein*)

**ITAM-** Motivos de ativação de imunoreceptores baseados em tirosinas (do inglês, *immunoreceptors tyrosine- based activation motifs*)

**ITIM-** Motivos de inibição de imunoreceptores baseados em tirosinas (do inglês, *immunoreceptors tyrosine-based inhibitory motifs*)

**IMAC** – Cromatografia de Afinidade por iões metálicos imobilizados (do inglês, *Immobilized-metal affinity chromatography*)

**IPTG** – Isopropil-β-D-tiogalactósido

**LB** – Luria-Bertani

**OD<sub>600nm</sub>** – Densidade ótica monitorizada a 600 nm

**PCR** – Reação em cadeia da polimerase (do inglês Polymerase Chain Reaction)

**PDB** – Protein Data Bank

**PTP-** Fosfatases de tirosinas (do inglês, *Protein Tyrosine Phosphatases*)

**Rpm** – Rotações por minuto

**RMN-** Ressonância Magnética Nuclear

**SAXS-** *Small angle X-ray scattering*

**SDS-PAGE** – Eletroforese em Gel de Poliacrilamida e Dodecil-Sulfato de Sódio (do inglês, *Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis*)

**SH2-** *Src homology-2*

**SHP-1-** Fosfatase com domínio de homologia Src-2-1 (do inglês, *Src homology-2 (SH2) domain-containing phosphatase-1*)

**TB-** *Terrific Broth*

**TEV** – *Tobacco Etch Virus nuclear-inclusion-a endopeptidase*

**TEMED-** *Tetramethylethylenediamine*

**Tir-** Receptor translocado de intimina (do inglês, *Translocated intimin Receptor*)

**Tris** – Tris(hidroximetil)aminometano

**Trx-** Tioredoxina

| Abreviaturas de aminoácidos |     |   |
|-----------------------------|-----|---|
| Alanina                     | Ala | A |
| Arginina                    | Arg | R |
| Asparagina                  | Asn | N |
| Aspartato                   | Asp | D |
| Cisteína                    | Cys | C |
| Glutamato                   | Glu | E |
| Glutamina                   | Gln | Q |
| Glicina                     | Gly | G |
| Histidina                   | His | H |
| Isoleucina                  | Ile | I |
| Leucina                     | Leu | L |
| Lisina                      | Lys | K |
| Metionina                   | Met | M |
| Fenilalanina                | Phe | F |
| Prolina                     | Pro | P |
| Serina                      | Ser | S |
| Treonina                    | Thr | T |
| Triptofano                  | Trp | W |
| Tirosina                    | Tyr | Y |
| Valina                      | Val | V |

## **1. Introdução**



## 1.1. Virulência bacteriana e controlo do hospedeiro

As infeções emergentes transmitidas por alimentos representam um grande problema para o sector de saúde pública, agrícola e alimentar.<sup>1,2</sup> Estas refletem a notável capacidade de certos agentes patogénicos microbianos (bactérias, vírus) em subverter os processos e a imunidade do hospedeiro. Para controlar as células hospedeiras, esses agentes patogénicos produzem um repertório de proteínas efetoras que têm como alvo fatores essenciais do hospedeiro, alterando o seu normal funcionamento, favorecendo a infeção. Entre os principais agentes patogénicos transmitidos aos humanos através de alimentos encontram-se duas estirpes de *Escherichia Coli*: a *Escherichia coli* enteropatogénica (EPEC) e *Escherichia coli* enterohemorrágica (EHEC).<sup>3</sup> Na mucosa intestinal, essas bactérias são conhecidas por provocarem lesões características designadas por *attaching and effacing* (A/E), uma vez que criam pedestais de actina, aos quais se ligam na superfície da célula hospedeira para destruir as microvilosidades intestinais.<sup>4,5</sup> Ambos os agentes patogénicos são responsáveis por doenças graves: EPEC é uma das principais causas de diarreia em crianças; enquanto que EHEC causa colite hemorrágica, trazendo complicações para a vida humana, sobretudo nos países em vias de desenvolvimento. Ainda não há tratamento específico contra estes patógenos.<sup>5,6</sup> O efector Recetor de intimina translocado (Tir) é um potencial alvo terapêutico, pois é essencial para o início da colonização intestinal por parte destas bactérias. Tir medeia a ligação bacteriana e ao mesmo tempo, subverte várias proteínas do hospedeiro.<sup>5,7</sup> Estudos estruturais detalhando as funções e mecanismos moleculares de Tir são, portanto, fundamentais para possíveis avanços terapêuticos.

Tir possui regiões que mimetizam sequências consenso presentes em proteínas do hospedeiro passíveis de sofrer fosforilação. Ao ser fosforilada, esta proteína efectora adquire a capacidade de recrutar e interagir com domínios específicos de proteínas alvo. De entre estas proteínas, destacam-se as fosfatases, nomeadamente a fosfatase SHP-1, uma proteína importante em algumas vias de sinalização da resposta imunitária. Ao interagir com SHP-1, Tir suprime a produção de citocinas pró-inflamatórias que inibem a multiplicação bacteriana.<sup>7</sup>

Nas secções seguintes, serão abordados alguns aspetos associados à patogenicidade bacteriana de dois agentes patogénicos (EPEC/EHEC) que utilizam o efector Tir para estabelecer infeções entéricas em humanos. Será de igual forma apresentada a estratégia mediada por Tir para estabelecer interações com proteínas alvo, como a SHP-1, de forma a bloquear a resposta imune e, consequentemente favorecer a infeção.

### 1.1.1. Patogenicidade bacteriana de *Escherichia coli* enteropatogénica e enterohemorrágica (EPEC/EHEC)

A *Escherichia coli* (*E. coli*) é uma bactéria gram-negativa descoberta em 1885. Desde então tem sido uma das espécies bacterianas mais estudadas e utilizada em estudos bioquímicos e

genéticos. Por ser um microrganismo de fácil cultivo e manipulação genética, tem também sido utilizada para a produção de proteínas recombinantes em laboratório.<sup>6,8</sup>

O habitat natural de *E. coli* é o trato gastrointestinal dos mamíferos e, em humanos, esta espécie representa o anaeróbio facultativo mais comum no microbioma intestinal.<sup>9</sup> Embora a maioria das suas estirpes sejam simbioses inofensivos, existem algumas estirpes patogênicas que podem causar uma variedade de doenças em animais e humanos.<sup>5,8</sup> Todas as estirpes de *E. coli* partilham um grupo central de genes que definem traços bioquímicos semelhantes. Contudo, algumas estirpes, nomeadamente as infecciosas, apresentam alguma diversidade genética, reflexo da adaptação comensal e patogénica, e da especificidade do hospedeiro. Essas divergências, estão presentes sobretudo nos genes virulentos, os quais são responsáveis pela distinção dos diferentes tipos de patógenos e da sua estirpe. Desta forma, cada patógeno codifica fatores de virulência característicos que permitem à bactéria causar infeções, com distintas implicações clínicas e patológicas.<sup>3,5</sup>

As estirpes de *E. coli* causadoras de infeções entéricas em humanos foram classificadas em seis patógenos distintos: *E. coli* enteropatogénica (EPEC), *E. coli* enterohemorrágica (EHEC), *E. coli* enterotoxigénica (ETEC), enteroagregativa (EAEC), *E. coli* enteroinvasiva (EIEC) e *E. coli* difusamente aderente (DAEC).<sup>5,10</sup>

A *Escherichia coli* enteropatogénica (EPEC) foi a primeira estirpe de *E. coli* a ser reconhecida como patogénica.<sup>6</sup> Este agente patogénico é um dos principais causadores de diarreia em humanos. Provoca diarreia prolongada, sendo esta normalmente acompanhada de febre leve e, por vezes, de vômitos.<sup>4,10,11</sup> EPEC tem a capacidade de infetar também animais, nomeadamente, ratos.<sup>12</sup> É um patógeno conhecido por ser uma das causas de mortalidade infantil em países em desenvolvimento, sobretudo em crianças com menos de dois anos.<sup>5</sup> À semelhança de EPEC, a *Escherichia coli* enterohemorrágica (EHEC) é também um patógeno zoonótico emergente. É responsável por causar gastroenterite aguda e colite hemorrágica em idosos e em crianças com menos de cinco anos.<sup>4,11,13</sup>

Apesar de nos últimos anos os números de infeção por EPEC e EHEC terem vindo a diminuir nos países desenvolvidos até ao ponto de já não serem consideradas um problema clínico, nos países em desenvolvimento, EPEC e EHEC continuam a ser uma das principais causas de infeções entéricas em crianças, constituindo ainda um sério problema de saúde pública e que para o qual ainda não existe um tratamento específico.<sup>5,14</sup>

A infeção por EPEC e EHEC está associada a uma histopatologia intestinal característica conhecida como *attaching and effacing* (A/E). O processo infeccioso inicia-se pela formação de pequenas microcolónias das células bacterianas à superfície da célula hospedeira, mais precisamente na membrana das células epiteliais do intestino (enterócitos). De seguida, a bactéria transloca um grupo de proteínas virulentas, chamadas de proteínas efetoras ou efetores através de um mecanismo de secreção de proteínas altamente especializado e conservado, o sistema de secreção do tipo III (T3SS, do inglês *Type III secretion system*). A canalização destes efetores para o interior do hospedeiro ocorre através de uma nanomáquina denominada por *injectisome* (arquitetura representada na figura 1.1A), fazendo com que as células bacterianas e

hospedeira consigam estabelecer um contacto direto, ou contacto íntimo, despoletando uma cascata de sinalização que culmina com a formação de pedestais de actina por baixo da região de adesão (fenómeno *attaching*).<sup>3,15-17</sup> Por sua vez, esta reorganização do citoesqueleto de actina permite à bactéria fixar-se firmemente à célula hospedeira, destruindo as microvilosidades do intestino (fenómeno *effacing*) (consultar a figura 1.1).<sup>3-5,18</sup>

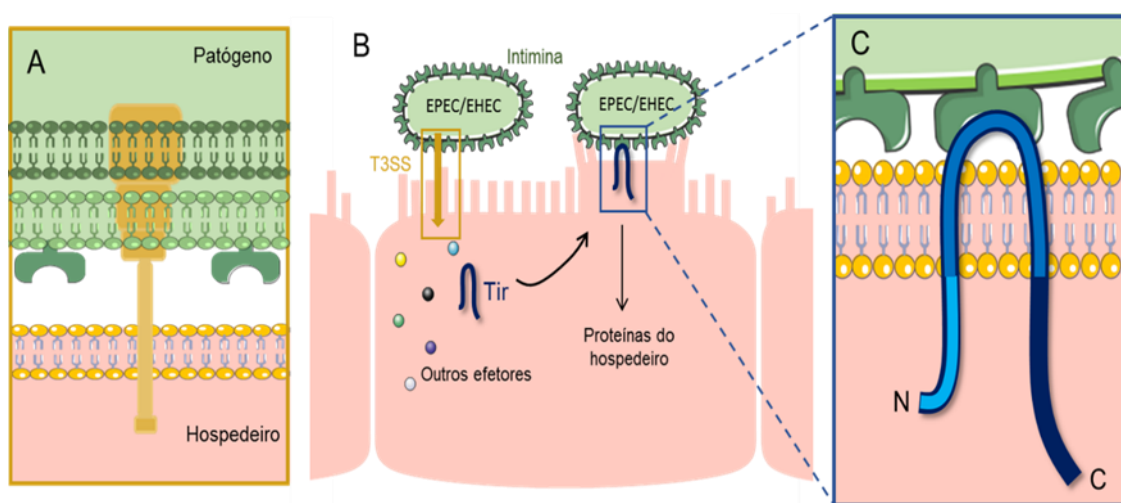
As lesões A/E são geneticamente codificadas por uma ilha de patogenicidade genómica de 35 kpb designada de *locus* de substituição de enterócitos (LEE, do inglês *Locus of Enterocyte Effacement*).<sup>19</sup> Esta região é considerada uma ilha de patogenicidade porque contém *loci* de virulência, que não são encontrados em estirpes não patogénicas de *E. coli*.<sup>3</sup> A região LEE codifica todos os componentes estruturais do sistema T3SS, sete efetores translocados através deste sistema e suas chaperonas, proteínas que participam na regulação da secreção e expressão dessa região LEE, bem como proteínas que estão envolvidas na aderência íntima da bactéria ao hospedeiro, como uma adesina bacteriana, a Intimina, e o seu recetor, Tir. Todos os genes presentes na região LEE de EPEC estão presentes na mesma ordem na região LEE de EHEC, com um elevado grau de conservação (94%).<sup>3,5,20,21</sup>

Este sistema de secreção canaliza até vinte e três proteínas efetoras bacterianas para o interior das células hospedeiras na infeção por EPEC e, até cinquenta por EHEC.<sup>22,23</sup> Geralmente, esses efetores apresentam sequências ou regiões semelhantes a nível estrutural e/ou funcional com as proteínas do hospedeiro. Essa é a estratégia utilizada pelos patógenos para colonizar os seus hospedeiros e ganhar o controlo de uma série de vias de sinalização celulares, garantindo a sua replicação e disseminação. Notavelmente, agentes patogénicos filogeneticamente distantes apresentam esse *modus operandi* na entrada e na subversão dos mecanismos da célula hospedeira, isto é, libertam efetores que possuem semelhanças moleculares com proteínas do hospedeiro, mimetizando as suas funções. Certos efetores mimetizam pequenos motivos lineares, que são pequenas regiões consenso de proteínas do hospedeiro, “fazendo-se passar” pelas mesmas, facilitando a sua internalização e manipulação celular.<sup>7,24-26</sup> De entre os efetores libertados por EPEC e EHEC, destaca-se o Recetor de intimina translocado (Tir), um efector que utiliza a estratégia mencionada para levar a cabo a infeção.

De seguida, será discutido em maior detalhe a forma como EPEC e EHEC translocam o efector Tir e como através da estratégia de mimetização molecular de certos motivos de proteínas do hospedeiro, conseguem interferir e ganhar o controlo de caminhos de sinalização celulares, promovendo a infeção.

### 1.1.2. Recetor de intimina translocado (Tir) – o mestre da virulência

Os patógenos EPEC e EHEC utilizam o T3SS para introduzir o seu próprio recetor na membrana plasmática das células do hospedeiro. Este recetor denominado por Recetor de intimina translocado (Tir, do inglês *Translocated Intimin Receptor*), uma proteína composta por ca 550 aminoácidos, que é libertada no citosol das células hospedeiras. Uma vez injetada no seu interior, este efector migra para a membrana plasmática, através de um processo desconhecido. Na membrana plasmática, Tir serve de recetor para uma adesina, a Intimina, um fator de virulência e proteína de membrana existente à superfície das células bacterianas.<sup>4,5,27,28</sup> Assim, EPEC e EHEC desenvolveram o seu mecanismo de aderência e contacto íntimo com a célula hospedeira, no qual as bactérias sintetizam a adesina (Intimina) e o seu recetor (Tir), sendo o último inserido diretamente na célula hospedeira.



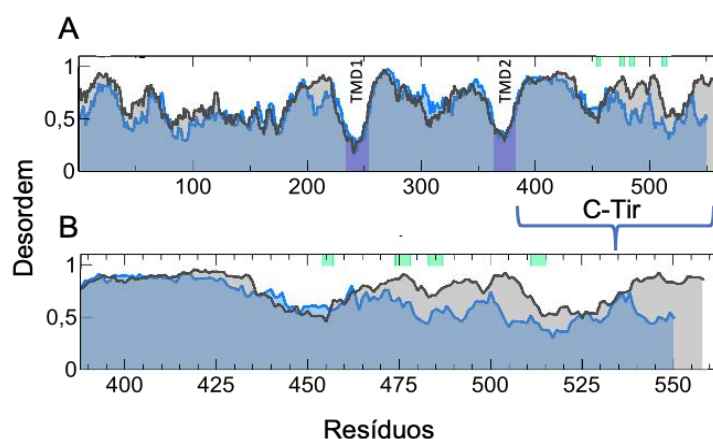
**Figura 1.1- Esquema representativo do modelo de patogenicidade de EPEC/EHEC mediada por Tir**  
Inicialmente EPEC/EHEC interage com as células intestinais de uma forma não íntima através do *injectisome* do T3SS. Este sistema transloca efetores essenciais à infeção para o interior da célula hospedeira. Destes efetores, destaca-se Tir que depois de injetado no citosol, migra para a membrana plasmática do hospedeiro, onde é promovido o contacto íntimo com a Intimina. Este contacto ativa de forma anormal cascatas de sinalização que resultam no recrutamento de proteínas do hospedeiro, para o ganho do controlo do mesmo promovendo a infeção. **(A)-** Representação do *injectisome* do T3SS. **(B)-** Modelo de aderência de EPEC/EHEC às células epiteliais do intestino via Tir e formação das lesões A/E. **(C)-** Topologia de Tir em forma de *hairpin*, inserida na membrana plasmática da célula hospedeira, ligada à Intimina e com os domínios N e C-terminais no interior do hospedeiro. Os diferentes tons de azul indicam as diferentes regiões de Tir, com destaque para C-Tir, representada a azul escuro.

Quando incorporada na membrana plasmática, Tir apresenta uma topologia em forma de *hairpin* com ambos os domínios N- e C-terminais localizados no citosol, ao passo que, o seu domínio extracelular está envolvido pela intimina, promovendo o contacto entre a bactéria e a célula hospedeira (figura 1.1B-C). Ao mesmo tempo que Tir expõe o seu domínio extracelular ao contacto, os seus domínios N- e C-terminais permanecem no citosol interagindo com proteínas do hospedeiro. Desta forma, estes domínios intracelulares fornecem um meio para o patógeno



estabelecer uma ligação direta ao interior do hospedeiro, iniciando cascatas de sinalização para alterar e ganhar o controlo dos seus mecanismos celulares.<sup>4,5,29</sup>

Tir demonstra uma tendência para desordem estrutural, conforme é visível na figura 1.2A, onde está apresentada a propensão de desordem estrutural de Tir-EPEC e EHEC.<sup>30</sup> Esta característica é enfatizada por este efector apresentar valores acima do limite de 0,5 de desordem estrutural, com destaque para a desordem da região C-terminal de Tir (C-Tir), apresentada na figura 1.2B. A flexibilidade estrutural presente em C-Tir permite que Tir tenha um mecanismo molecular capaz de realizar múltiplas interações, funcionando ao mesmo tempo como um local de ancoragem para várias proteínas do hospedeiro e como um fator de sinalização intracelular. De facto, esta proteína está presente numa vasta rede de interações regulatórias, interagindo com pelo menos vinte e cinco proteínas do hospedeiro.<sup>31</sup> Muitas destas proteínas, como SHP-1, SHP-2, SHIP-1, SHIP-2, Nck, IRSp53, entre outras, incluem domínios característicos como domínios SH2 e SH3 e estão implicadas na organização do citoesqueleto e na remodelação da membrana.



**Figura 1.2- Propensão da desordem estrutural de Tir dos agentes patogénicos EPEC e EHEC.** A azul e preto estão apresentadas as análises das propensões de desordem de Tir de EPEC e EHEC, respetivamente, calculadas através do *software* IUPred2A.<sup>30</sup> A verde estão representadas quatro tirosinas, compreendidas em quatro regiões mimetizadas por Tir-EPEC, passíveis de sofrer fosforilação por cinases do hospedeiro. A área a roxo representa os domínios transmembranares (TMD1 e TMD2). **(A)**- Propensão da desordem estrutural de Tir de EPEC e EHEC. **(B)**- Propensão da desordem estrutural da região C-terminal de Tir (C-Tir) de EPEC e EHEC.

A presente dissertação visou o estudo da região C-terminal de Tir (C-Tir) da estirpe EPEC. Além das previsões calculadas, anteriormente no laboratório já foi verificado que C-Tir se comporta como uma proteína intrinsecamente desordenada (IDP, do inglês *Intrinsically Disordered Protein*) (Vieira M. *et al.* artigo em preparação). As características estruturais deste tipo de proteínas encontram-se apresentadas de forma detalhada mais à frente no documento (consultar secção 1.2.1.).

De entre as interações estabelecidas por Tir referidas anteriormente, destacam-se as que reportam para as vias de sinalização que este efector inicia na célula hospedeira que culminam com a formação dos pedestais de actina essenciais à adesão bacteriana. Também as interações com os domínios específicos referidos de proteínas alvo têm particular importância, pois graças a estas, a resposta imune é bloqueada e a infecção é favorecida. Para isso, são fundamentais quatro regiões consenso de proteínas eucariotas do hospedeiro que C-Tir mimetiza e possui na sua composição, que contêm resíduos de tirosina passíveis de sofrer fosforilação por cinases do hospedeiro (resíduos representados a verde na figura 1.2).

Na secção seguinte, será dado destaque a essas regiões consenso mimetizadas por C-Tir, importantes para o ganho do controlo dos mecanismos de sinalização celular do hospedeiro, conduzindo-o à infecção.

#### **1.1.2.1. Interação entre Tir e proteínas com domínios SH2 – estratégia de infecção e inibição de resposta imune**

Nos seres eucariontes, as respostas imunitárias são geralmente reguladas positivamente através de recetores que contêm motivos de ativação de imunorecetores baseados em tirosinas (ITAMs, do inglês *immunoreceptors tyrosine-based activation motifs*). Quando os ligandos ativam os recetores, as tirosinas presentes nestes ITAMs são rapidamente fosforiladas. Uma vez fosforilados, estes recetores formam locais de ancoragem para a ligação de outras proteínas celulares que contêm domínios *src-homology 2* (SH2) na sua composição, como proteínas cinases de tirosinas, e outras moléculas adaptadoras que traduzem uma cascata de sinalização, fosforilando outras moléculas efectoras essenciais à imunidade do organismo. Algumas das respostas imunitárias que são mediadas por ITAMs incluem a fagocitose, produção de citocinas e a apresentação de antígenos.<sup>26,32</sup>

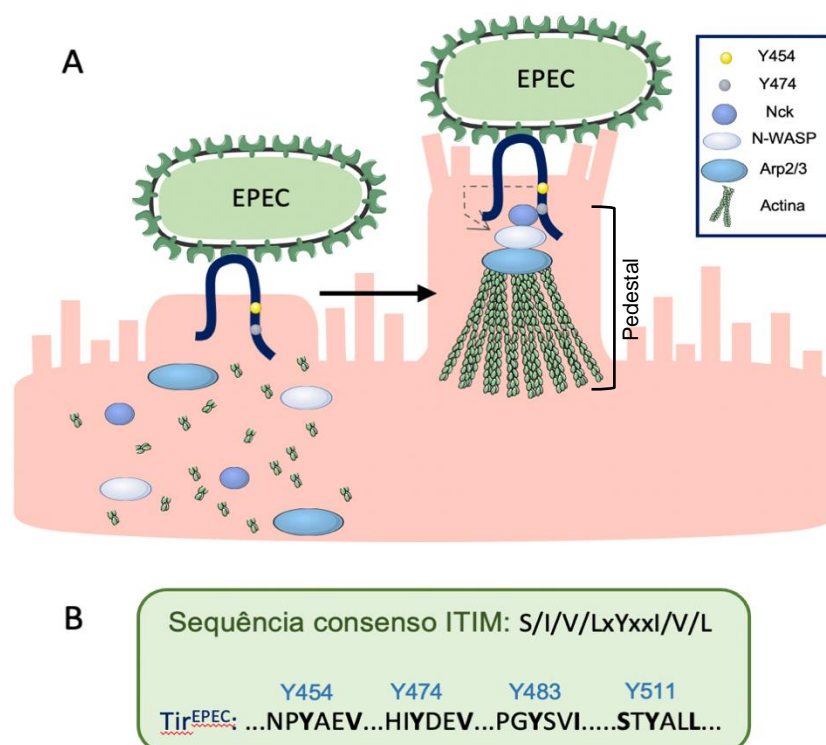
Por outro lado, as respostas imunitárias são reguladas negativamente através de recetores que contêm motivos de inibição de imunorecetores baseados em tirosinas (ITIMs, do inglês *immunoreceptors tyrosine-based inhibitory motifs*). O mecanismo inibitório depende, à semelhança do que acontece com os ITAMs, da fosforilação das tirosinas por cinases intrínsecas do hospedeiro. Estas cinases são normalmente cinases da família Src, como é o caso da cinase Fyn. De igual forma, esta fosforilação forma locais de ligação para proteínas com domínios SH2.<sup>7,26</sup>

Tal como referido anteriormente, alguns agentes patogénicos desenvolveram a estratégia de mimetização molecular para levarem a cabo a infecção. Tir é um efector essencial translocado por EPEC e EHEC, que notavelmente, apresenta na sua região C-terminal (C-Tir), quatro regiões de similaridade para as regiões consenso dos ITIMs eucariotas (S/I/V/LxYxxI/V/L) (figura 1.3B). Um outro exemplo do descrito, é a proteína CagA produzida e secretada pela bactéria *Helicobacter pylori*, um agente causante de gastrite que pode evoluir para cancro.<sup>33</sup> Desta maneira, Tir mimetiza esses ITIMs do hospedeiro eucariótico, para mediar o recrutamento de proteínas com domínios SH2, e consequentemente, estabelecer a infecção.<sup>7,26</sup>

O citoesqueleto de actina das células é vital para uma variedade de processos celulares, incluindo a migração, proliferação e apoptose celular.<sup>34</sup> EPEC e EHEC, através da translocação de Tir, apresentam a capacidade para subverterem o mecanismo de actina do hospedeiro, de forma a promover a sua aderência e disseminação e, conseqüentemente, favorecer a infecção. Estes patógenos reorganizam o citoesqueleto do hospedeiro formando os pseudópodes de F-actina também denominados pedestais de actina, por baixo da bactéria anexada à célula hospedeira. Desta forma, é promovida a colonização bacteriana e a fixação permanente da bactéria à célula hospedeira, provocando a destruição das microvilosidades do intestino e a posterior doença diarreica.<sup>5,34,35</sup>

Para a formação destes pedestais, é essencial a interação referida anteriormente entre Tir e a Intimina. No caso de EPEC, esta interação aciona eventos de sinalização como a fosforilação das tirosinas de dois dos ITIMs mimetizados por Tir (Y454 e Y474). As cinases da família-Src, nomeadamente a cinase Fyn é a responsável pela fosforilação do resíduo de tirosina Y474.<sup>34,36</sup> Quando fosforilado, este resíduo recruta o domínio SH2 da proteína adaptadora Nck do hospedeiro.<sup>37</sup> Ao interagir com a tirosina Y474, Nck ativa a proteína da síndrome neural Wiskott-Aldrich (N-WASP), que por sua vez estimula a ativação do complexo proteico Arp2/3 levando à polimerização de actina e à formação dos pedestais.<sup>15,34,38</sup> Por outro lado, a fosforilação da tirosina Y454 desencadeia um caminho independente da proteína adaptadora Nck, que é menos eficiente, mas que termina da mesma maneira: o recrutamento de N-WASP e com a formação dos pedestais de actina (figura 1.3A).<sup>35</sup>

Curiosamente, em EHEC a formação dos pedestais de actina é independente da fosforilação da tirosina Y474 e da proteína adaptadora Nck, uma vez que esse resíduo não existe. Em vez disso, a acumulação de actina é conseguida através de outro efector T3SS, o EspF<sub>U</sub>. Este recruta N-WASP, que por sua vez ativa o complexo proteico Arp2/3, levando à polimerização de actina.<sup>15,35</sup> A interação entre Tir-EHEC e EspF<sub>U</sub> é conseguida através de outra proteína, a IRSp53. Esta interação ocorre através do motivo NPY de Tir, que compreende a tirosina Y458 (correspondente à tirosina Y454 em Tir EPEC) e o domínio I-BAR de IRSp53. Este motivo NPY é reconhecido por interagir com proteínas com domínios I-BAR.<sup>39,40</sup> Uma vez que EPEC e EHEC têm sequências conservadas e ambas apresentam o motivo NPY, espera-se que o mesmo possa acontecer com EPEC, embora a interação referida ainda não tenha sido demonstrada.



**Figura 1.3- Mimetização de regiões ITIM e controlo do hospedeiro por Tir EPEC. (A)-** Representação esquemática da formação dos pedestais de actina por EPEC. Quando Tir-EPEC se encontra inserida na membrana plasmática ligada à intimina, sofre uma fosforilação no resíduo Y474 por cinases da família Src, desencadeando o recrutamento da proteína adaptadora Nck do hospedeiro, que interage com o resíduo de tirosina através do seu domínio SH2. Esta liga-se à N-WASP que por sua vez estimula a ativação do complexo proteico Arp2/3, levando à polimerização da actina e consequente formação dos pedestais de actina. EPEC também pode gerar os pedestais por meio de caminhos independentes de Nck, mas com menos eficiência (linha cinzenta a tracejado). **(B)-** Representação dos ITIMs eucariotas mimetizados por Tir, responsáveis por mediar o recrutamento de domínios SH2 de proteínas do hospedeiro, conduzindo à infecção. Os resíduos que correspondem à sequência consenso de ITIMs celulares estão indicados a negrito.

Os *Toll-like receptors* (TLRs) são recetores que estão presentes nas membranas celulares do hospedeiro e que têm um papel central na iniciação da resposta imune contra os agentes patogénicos.<sup>41</sup> Estes recetores reconhecem moléculas associadas à patogenicidade bacteriana, como é o caso dos lipopolissacarídeos.<sup>42</sup> Quando os reconhecem, recrutam proteínas adaptadoras, como TRAF6, ativando múltiplas cascatas de sinalização intracelular e a biossíntese de moléculas imunorreguladoras. É conhecido que a interação entre Tir e SHP-1 facilita o recrutamento desta fosfatase para a proteína adaptadora TRAF6, inibindo as suas funções. Desta forma, a mimetização de quatro regiões consenso de ITIMs eucariotas por parte de Tir, fazendo com que a SHP-1 seja recrutada é o ponto chave do controlo da resposta imunitária que Tir assume com o hospedeiro.<sup>7,43</sup>

Embora a interação entre Tir e SHP-1 já tenha sido demonstrada e documentada anteriormente, as dinâmicas moleculares, a afinidade e a estequiometria desta interação ainda

não são conhecidas.<sup>7</sup> Assim, a compreensão desta interação assume um papel crucial para o entendimento da forma como Tir regula indiretamente a resposta imunitária e, consequentemente, favorece a infecção.

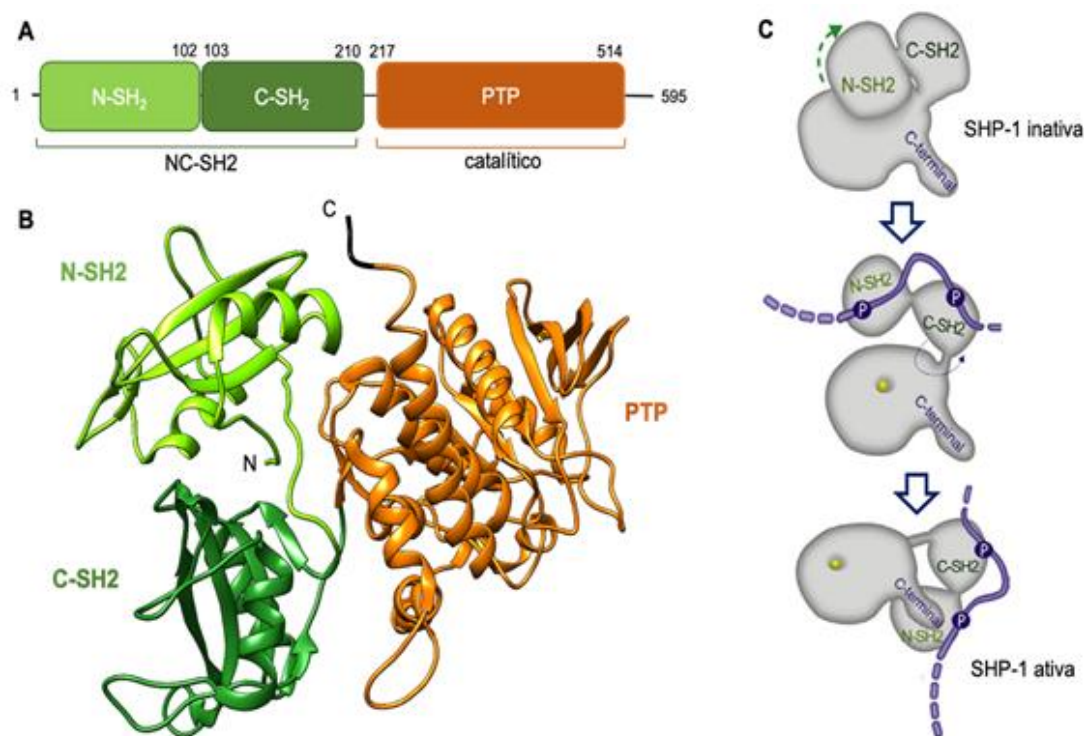
### 1.1.3. Fosfatase com domínio de homologia Src-2-1 (SHP-1)

As fosfatases com domínios de homologia Src-2 (SHPs, do inglês *Src homology-2 (SH2) domain-containing phosphatases*) são uma subfamília das proteínas fosfatases de tirosinas (PTP, do inglês *Protein Tyrosine Phosphatases*) não transmembranares. Esta subfamília é composta por duas proteínas: a SHP-1 e SHP-2. Apesar de haver uma similaridade geral e estrutural entre SHP-1 e SHP-2, as suas funções e a forma como se comportam são distintas.<sup>44,45</sup> O facto de ambas as fosfatases conterem dois tipos de domínios capazes de interagir com resíduos de tirosina fosforilados e, por sua vez, participar em diversos caminhos de sinalização celulares, fez com que o estudo das SHPs tivesse suscitado grande interesse na comunidade científica, sobretudo na forma como estas se comportam a nível estrutural para regular cataliticamente algumas das vias de sinalização celular.<sup>44,46</sup>

A fosfatase abordada nesta dissertação é a SHP-1. Esta fosfatase é expressa maioritariamente em células hematopoiéticas, embora também seja expressa em algumas células epiteliais, em astrócitos e oligodendrócitos do sistema nervoso central, ainda que nestes últimos em muito menor quantidade.<sup>47-49</sup>

Embora os mecanismos concretos de ação de SHP-1 precisarem de ser elucidados, esta proteína parece mediar a inibição das etapas iniciais das vias de sinalização ativadoras. Isto é, está implicada na regulação negativa de uma gama diversificada de vias de sinalização ativadoras. Entre estas, a via de sinalização dos leucócitos, nomeadamente dos linfócitos T ou células T, é a que vem sido descrita como a via onde SHP-1 tem maior importância, tendo um papel fundamental na regulação destas macromoléculas, envolvidas na maioria das respostas imunitárias.<sup>47,50,51</sup>

Estruturalmente, SHP-1 contém dois domínios SH2 na região N-terminal, os domínios *amino* (N) e *carboxi* (C) SH2 (N-SH2 e C-SH2), um domínio catalítico clássico PTP desta família de proteínas e uma zona de aminoácidos básicos na região C-terminal, denominada de *C-tail* (figura 1.4A-B).



**Figura 1.4- Estrutura geral de SHP-1.** O domínio N-SH2 está representado a verde claro, o domínio C-SH2 a verde escuro e o domínio catalítico PTP a cor de laranja. **(A)**-Organização dos domínios de SHP-1. **(B)**-Estrutura tridimensional de SHP-1 (código PDB: 3ps5). **(C)**- Mecanismo de ação de SHP-1. No estado basal, a fosfatase permanece inativa pois o domínio N-SH2 bloqueia estereoquimicamente o domínio catalítico PTP. O domínio C-SH2 ao interagir com um ligando induz uma alteração conformacional na fosfatase permitindo a abertura do domínio catalítico PTP, permitindo o acesso deste aos seus substratos. Esquema adaptado de <sup>46</sup>.

Tal como referido anteriormente, os domínios SH2 têm a capacidade de se ligarem com elevada afinidade a sequências específicas que contêm resíduos de tirosina fosforilados. É dessa forma que SHP-1 é recrutada pelos seus ligandos.<sup>47</sup> A ligação a esses resíduos fosforilados ocorre à superfície de SHP-1, longe do seu domínio PTP e a orientação de cada domínio SH2 em relação ao domínio PTP é diferente. Enquanto N-SH2 estabelece contactos com o domínio catalítico, C-SH2 não os estabelece, sugerindo ter mais mobilidade e estar mais suscetível a ligar-se aos substratos que compreendem tirosinas fosforiladas.<sup>47,52</sup>

Esta proteína tem duas conformações possíveis: a conformação inativa (quando não tem capacidade para interagir com os ligandos e não exerce a sua função sob os seus substratos, por restringimentos estereoquímicos) e a conformação ativa (quando não há impedimento estereoquímico, possibilitando a interação com os ligandos e o recrutamento dos seus substratos).

No seu estado basal, SHP-1 não está associada a nenhum ligando, pois, o domínio N-SH2 estende-se até ao domínio catalítico PTP, obstruindo-o, mantendo a fosfatase na sua conformação inativa. A interação do domínio C-SH2 com um ligando induz uma alteração

conformacional nesta macromolécula, na região à superfície de N-SH2. Este rearranjo conformacional em N-SH2, permite a consequente abertura do domínio PTP, tornando suscetível o acesso de SHP-1 aos seus substratos. Desta maneira, é previsível que o domínio N-SH2 exista num equilíbrio dinâmico entre as duas conformações como um interruptor alostérico regulado potencialmente pelo domínio C-SH2 (figura 1.4C).<sup>46,47</sup> Assim sendo, na maioria das vezes, o domínio C-SH2 é quem inicia a interação com o ligando que contém as tirosinas fosforiladas.

Os estudos estruturais foram essenciais na identificação das macromoléculas que interagem com SHP-1. Desde aí, foram descobertas muitas macromoléculas que interagem com esta fosfatase, como diferentes antígenos, citocinas, recetores de fatores de crescimento, adaptadores *scaffolding*, entre outras.<sup>47</sup>

As evidências estruturais e bioquímicas atuais apontam para que SHP-1 seja recrutada preferencialmente por moléculas portadoras da sequência consenso e de referência de ITIMs de eucariontes (I/V/LxY(p)xxI/V/L). O recrutamento e consequente ativação de SHP-1 só acontece quando a tirosina presente na sequência consenso referida se encontra fosforilada.<sup>44,46,47,53,54</sup> Esta evidência é facilmente compreendida, uma vez que SHP-1 é uma fosfatase da PTP, com especificidade para reconhecer resíduos de tirosina fosforilados.

A ligação de ambos os domínios SH2 (N-SH2 e C-SH2) às regiões ITIM estão associadas a uma maior atividade catalítica de SHP-1. Da mesma maneira, os estudos realizados com esse propósito sugerem que, na maioria dos casos, a ativação da fosfatase *in vivo* necessita mesmo da ligação de ambos os seus domínios SH2 ao substrato. Desta maneira, os substratos que possuam duas ou mais sequências consenso de ITIMs suscetíveis de sofrer fosforilações, e que, ao mesmo tempo, sejam capazes de interagir com domínios SH2, apresentam-se como os mais relevantes do ponto de vista fisiológico.<sup>46,47</sup> Neste sentido, a região C-terminal desordenada de Tir possui na sua composição quatro destas sequências, tornando-se o substrato perfeito para o recrutamento de SHP-1.

## **1.2. Estrutura, dinâmica e interação de Proteínas Intrinsecamente Desordenadas (IDPs)**

Durante muito tempo, existiu o paradigma que postulava que a funcionalidade específica de uma proteína era predeterminada pela sua estrutura tridimensional. De acordo com esse conceito, uma proteína adquire a sua função biológica, quando se enrola até atingir o seu estado estruturado, representativo da conformação energeticamente mais favorável. O estado em que uma proteína perde a sua função, ou estado desnaturado, estaria apenas associado à perda da sua estrutura tridimensional, por estar sujeita a condições de elevada temperatura ou pressão, variações extremas de pH ou na presença de detergentes ou agentes caotrópicos. Contudo, ao longo dos últimos anos, foram desenvolvidos estudos em proteínas, que demonstraram ser contraditórios a esse paradigma. Foram descobertas proteínas que, apesar de não possuírem

uma estrutura tridimensional ordenada e definida, desempenhavam papéis biológicos muito importantes. Este tipo de proteínas são, hoje em dia, amplamente conhecidas como proteínas intrinsecamente desordenadas (IDPs, do inglês *Intrinsically Disordered Proteins*). Um exemplo de uma IDP é o efector virulento abordado nesta dissertação: Tir.

### 1.2.1. IDPs

As proteínas intrinsecamente desordenadas (IDPs) são uma classe de proteínas que não têm estrutura tridimensional bem definida, sob condições fisiológicas. Este tipo de proteínas compõem um terço do proteoma humano.<sup>55,56</sup> A sua descoberta e caracterização tem evoluído imenso, alargando a visão da comunidade científica acerca do dogma central da biologia estrutural, tornando-se mesmo numa das áreas da proteómica em maior crescimento.<sup>57</sup>

Este dogma central afirma que a estrutura tridimensional de uma proteína determina a sua função e é absolutamente necessária para a mesma. No entanto, a visão determinística sobre a função das proteínas tem sido desafiada pelo campo em expansão de proteínas intrinsecamente desordenadas. Atualmente, é aceite pela comunidade científica que o destino funcional de uma dada cadeia polipeptídica é determinado pelas particularidades dos aminoácidos da sua sequência e não pela sua estrutura, por si só.<sup>57</sup> No geral, as IDPs exibem baixa complexidade sequencial, apresentando um número baixo de resíduos hidrofóbicos volumosos (Val, Leu, Ile, Met, Phe, Trp, Tyr), os quais estão presentes normalmente no *core* das proteínas ordenadas. Contrariamente, apresentam um número considerável de resíduos polares e carregados (Gln, Ser, Pro, Glu, Lys, Gly, Ala).<sup>58,59</sup>

Ainda que muitas proteínas sejam ordenadas em estruturas tridimensionais únicas, possuindo funções biológicas concretas, estas tendem a desenrolar-se espontaneamente. A flexibilidade estrutural e a plasticidade, representam a maior vantagem funcional das IDPs. Essas características permitem que este tipo de proteínas superem as proteínas ordenadas em certas funções. As IDPs não só toleram esse estado estrutural desordenado, como é essa desordem que as predispõe a modos funcionais especiais, os quais tiram vantagem disso mesmo. Para compreender os papéis da desordem, é importante considerar as vantagens que a flexibilidade oferece à medida que essas proteínas evoluem para funções específicas onde a flexibilidade conformacional é uma vantagem. Essa flexibilidade e a capacidade dessas proteínas flutuarem ao longo de um conjunto de estados estruturais permitem que as mesmas interajam com vários parceiros, apresentando o que se chama de promiscuidade de ligação. Como tal, estas proteínas são frequentemente encontradas no centro de diversas redes de interação de proteínas, estando envolvidas em interações regulatórias e de sinalização com múltiplas proteínas, sendo parte integrante de processos como a regulação da divisão celular, a transcrição e tradução, a tradução de sinal, a ação de chaperonas e a regulação de grandes complexos multiproteicos. Além disso, as IDPs são importantes alvos de processos de regulação adicionais governados por modificações pós-traducionais, como a fosforilação.<sup>60-64</sup>

Um aspeto fundamental para explicar a associação da desordem a fenómenos de sinalização e regulação está relacionado com o modo como estas proteínas reconhecem os seus parceiros,



pois apesar destas serem desordenadas em solução, desempenham as suas funções biológicas vinculando-se aos mesmos. As interações das IDPs com os seus parceiros envolvem, geralmente, uma transição a nível estrutural, passando de uma conformação desordenada a ordenada ou parcialmente ordenada. Estas transições têm um custo entrópico, o que implica muitas vezes afinidades baixas a moderadas.<sup>56</sup> Essa particularidade em combinação com a alta especificidade das interações torna as IDPs agentes muito adequados para processos de sinalização e de regulação.<sup>60,61,64</sup> Da mesma maneira, existem alguns exemplos de interações onde não se verifica nenhuma estruturação após a ligação e permanecendo a proteína desordenada no complexo. Estas situações acontecem geralmente quando a interação é mediada por forças electrostáticas.<sup>60,61,64</sup>

Dado o número de processos biológicos importantes mencionados acima que estão relacionados com a desordem proteica, não é surpreendente que a maioria das proteínas envolvidas em doenças como cancro, doenças cardiovasculares ou neurodegenerativas sejam predominantemente desordenadas.<sup>56,65</sup>

Como referido, a plasticidade conformacional, a flexibilidade e a estruturação transiente são características inerentes às IDPs e que as tornam as candidatas ideais a modular o reconhecimento dos seus parceiros em processos de sinalização. Como consequência, a caracterização das suas propriedades estruturais, tanto no seu estado livre, como em complexo com os seus parceiros biológicos, é crucial para revelar a base molecular da sinalização e da regulação celular.<sup>56,66</sup> Existem alguns métodos biofísicos capazes de dar resposta a essas questões. Exemplos desses métodos são: o CD, o SAXS e a espectroscopia de RMN. Enquanto que CD é sensível aos elementos de estrutura secundária, o SAXS é extensivamente utilizado no estudo destas proteínas flexíveis a baixa resolução. Contudo, a espectroscopia de RMN é a única técnica capaz de fornecer informação estrutural e dinâmica a nível atómico das IDPs. Para além disso, a sensibilidade da técnica de espectroscopia de RMN ao ambiente químico permite a monitorização de padrões de reconhecimento e de modificações pós-traducionais, como a fosforilação.<sup>56</sup> Desta forma, a técnica utilizada ao longo deste trabalho, com o intuito de seguir a fosforilação da IDP de interesse (Tir) e a sua interação com os domínios SH2 de SHP-1, foi a espectroscopia de RMN.

### **1.2.2. Ressonância Magnética Nuclear em IDPs**

A espectroscopia de Ressonância Magnética Nuclear (RMN) desempenha um papel muito importante na caracterização de IDPs, pois é a única técnica espectroscópica capaz de fornecer informações estruturais e dinâmicas a nível atómico de proteínas desordenadas e altamente dinâmicas.<sup>56,67</sup>

Para realizar esses estudos, é essencial a atribuição das ressonâncias dos desvios químicos. Contudo, algumas características das IDPs provocam consequências nos espectros, dificultando esse processo. Uma dessas consequências é a baixa dispersão nos desvios químicos dos prótons das amidas. Devido à elevada dinâmica e flexibilidade das IDPs, os seus desvios

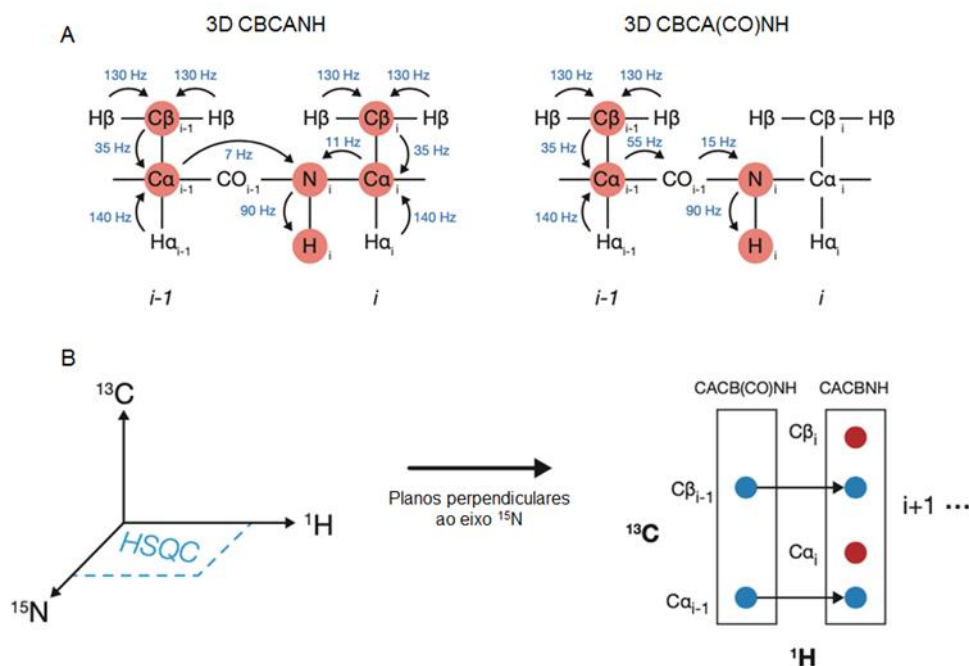
químicos colapsam na região dos desvios químicos das conformações *random-coil* (a cerca de  $^1\text{H}$  7.5-8.5 ppm), causando sobreposições de ressonâncias. As IDPs ao se encontrarem em constante movimento, estão muito expostas ao solvente, permitindo extensos processos de permuta com o mesmo. Desta maneira, como os desvios químicos são sensíveis ao ambiente químico local dos núcleos, as ressonâncias dos seus desvios químicos ocorrem a frequências quase idênticas, dando origem a espectros com picos sobrepostos. No passado, a atribuição dos desvios químicos nessas condições era um desafio muitas vezes inultrapassável. No entanto, a disponibilidade de espectrômetros de campo magnético mais alto e a possibilidade de registrar espectros de elevada resolução num tempo razoável, graças à *non-uniform sampling*, levaram a uma melhoria significativa da capacidade de resolver esse desafio.<sup>56,68</sup>

A exposição das IDPs ao solvente causa também outra consequência nos espectros de RMN: o alargamento dos sinais. Devido à falta de elementos estruturados, as IDPs estão muito expostas ao solvente, em comparação com as proteínas com estrutura tridimensional. Essa exposição leva a extensos processos de permuta entre os prótons da amida e os da água, resultando na perda de intensidade dos sinais de correlação  $^1\text{H}$ - $^{15}\text{N}$ , o que consequentemente, resulta numa diminuição de sensibilidade e resolução dos espectros.<sup>67,69</sup> Felizmente, as IDPs são geralmente estáveis num intervalo grande de pH, o que permite trabalhar a um pH ligeiramente ácido, de forma a reduzir estas permutas. Além disso, também a redução da temperatura pode ser usada para diminuir estas trocas e adquirir espectros com melhor qualidade.<sup>56,70</sup>

Alternativamente, no caso de proteínas com maior peso molecular, uma abordagem apelidada de “dividir para conquistar” é frequentemente usada, para reduzir a complexidade dos espectros. Um exemplo dessa abordagem é Tir, onde no laboratório em que foi realizada esta dissertação se está a trabalhar nas regiões N- e C-terminais isoladamente, com o objetivo final de obter a atribuição das ressonâncias de toda a sequência.

Geralmente, a assinatura ou *fingerprint* de uma proteína é dada pela experiência de bidimensional heteronuclear  $^1\text{H}$ - $^{15}\text{N}$  HSQC. Esta correlaciona o azoto (N) do grupo amida com o próton (H) que lhe está diretamente ligado, quer sejam estes da cadeia principal ou da cadeia lateral dos aminoácidos. Com base nesta experiência 2D, são realizadas uma série de experiências multidimensionais para atribuir as ressonâncias de RMN à sequência de uma proteína. Nessas experiências, em vez de se considerar apenas o acoplamento  $^1J_{\text{NH}}$  (correlações NH referidas), são explorados vários acoplamentos J para atribuir as frequências dos átomos da estrutura da proteína, estabelecendo correlações espectrais entre os átomos dos planos peptídicos contíguos entre si, permitindo uma atribuição sequencial.<sup>67,71</sup>

Duas experiências multidimensionais que são frequentemente usadas com esse propósito são as experiências 3D CBCANH e CBCA(CO)NH. Conforme é observável na figura 1.5A, a primeira destas experiências correlaciona o HN e o N de um resíduo  $i$  com os C $\alpha$  e C $\beta$  do próprio resíduo e do resíduo anterior  $i-1$ . Desta maneira, o espectro desta experiência contém correlações inter e intraresiduais. Por outro lado, a segunda experiência referida, correlaciona o C $\alpha$  e C $\beta$  do resíduo anterior  $i-1$  com o NH e o N do próprio resíduo ( $i$ ).<sup>71-73</sup>



**Figura 1.5- Atribuição das ressonâncias da cadeia principal de uma proteína com recurso a experiências 3D de RMN. (A)-** Experiências 3D tipicamente realizadas para a atribuição sequencial dos sistemas de spin de uma proteína. O  $i-1$  representa o resíduo anterior ao resíduo  $i$ . **(B)-** Representação esquemática de um par de experiências 3D de RMN realizadas de forma a atribuir as ressonâncias da cadeia principal de uma proteína.

Extrapolando do espectro HSQC temos as experiências 3D representadas num cubo, onde o plano xy contém a frequência das correlações NH e a terceira dimensão corresponde à frequência dos carbonos (figura 1.5B). A extração de planos perpendiculares ao eixo  $^{15}\text{N}$  permite a conexão dos resíduos  $i$  com os seus resíduos adjacentes  $i-1$ . Isto associado à informação dos padrões de dispersão específicos dos desvios químicos dos carbonos, permite a atribuição de uma maneira específica da sequência da proteína num espectro HSQC, permitindo avançar com os estudos de espectroscopia de RMN subsequentes.<sup>74</sup>

Uma vez conseguida a atribuição das ressonâncias dos desvios químicos da proteína, existem vários parâmetros observáveis de RMN que podem ser utilizados para retirar informação dinâmica e estrutural. Alguns exemplos desses parâmetros são os desvios químicos, acoplamentos escalares, acoplamentos dipolares residuais (RDCs, do inglês *Residual Dipolar*

*Couplings*), padrões de relaxação (R1 e R2), bem como o efeito do aumento de relaxação paramagnético (PRE, do inglês *Paramagnetic Relaxation Enhancement*).<sup>56,75</sup>

No âmbito deste trabalho, apenas tiramos partido da informação contida nos desvios químicos. Este é o parâmetro mais acessível num espectro de RMN e, dada a sua sensibilidade ao ambiente químico do núcleo, torna-o num excelente indicador da conformação estrutural de uma proteína. As alterações dos desvios químicos medidos experimentalmente em relação a valores dos mesmos em conformação *random coil* são usados para estudar a propensão de estrutura secundária.<sup>76</sup> Essas alterações são designadas por desvios químicos secundários (em inglês, *secondary chemical shifts*) e são definidas por  $\Delta\delta = \delta - \delta_{RC}$ , em que  $\delta$  é o desvio químico observado e  $\delta_{RC}$  é o desvio químico em conformação *random coil*. Valores positivos de desvios químicos secundários do C $\alpha$  ( $\Delta\delta^{C\alpha}$ ) indicam propensão de estruturas helicais, enquanto que valores negativos indicam a preferência para folhas- $\beta$ . Para os valores de desvios químicos secundários do C $\beta$  ( $\Delta\delta^{C\beta}$ ) o efeito é o inverso e, portanto, é conveniente monitorizar a diferença entre  $\Delta\delta^{C\alpha}$  e  $\Delta\delta^{C\beta}$ . Dessa forma, valores positivos da diferença  $\Delta\delta^{C\alpha} - \Delta\delta^{C\beta}$  indicam tendência para elementos de estrutura secundária de hélice- $\alpha$  e valores negativos indicam preponderância para folhas- $\beta$ .<sup>56,67,77–80</sup>

A partir da definição de  $\Delta\delta$ , é fácil perceber que o resultado das previsões da estrutura secundária com base em medidas de desvios químicos secundários depende em grande parte dos desvios químicos da conformação *random coil*. A abordagem tradicional baseia-se no uso de valores de  $\delta_{RC}$  genéricos determinados. A precisão dos  $\Delta\delta$  depende, portanto, da qualidade desses valores. Este aspeto torna-se ainda mais crucial nas IDPs, uma vez que os desvios químicos medidos experimentalmente são geralmente próximos dos valores dos desvios químicos *random coil* correspondentes. Dessa forma, o cálculo dos  $\Delta\delta$  para IDPs deve ser realizado utilizando desvios químicos em conformação *random coil* corrigidos para este tipo de proteínas.<sup>56,78,81,82</sup>

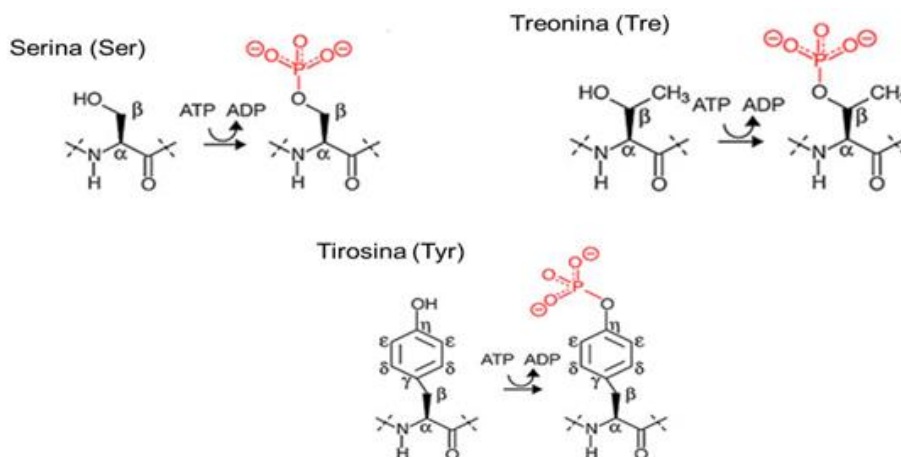
Os  $\Delta\delta$  podem ser calculados considerando os resíduos individualmente. Contudo, para melhorar a eficiência e a confiabilidade das previsões de propensão de estrutura secundária, pode ser utilizado o método *Neighbor Corrected Structural Propensity Calculator* (ncSPC), que considera a influência dos resíduos vizinhos de cada resíduo.<sup>82</sup>

Os desvios químicos são também uma ferramenta importante para estudar modificações pós-traducionais. Esses fenómenos participam frequentemente na regulação das interações estabelecidas pelas IDPs. Este tipo de proteínas, em parte graças à sua desordem estrutural, são alvo constante destas modificações, nomeadamente de fosforilações. A fosforilação é a modificação pós-traducional mais comum estudada através de RMN. Esta técnica permite não só identificar a presença dessa modificação na proteína e localizá-la ao nível da sua sequência, mas também seguir a sua cinética e as mudanças estruturais por ela causada. Esta capacidade é fundamental para compreender o papel dessa modificação no contexto de reconhecimento e interação de parceiros biológicos.<sup>56,83</sup>

### 1.2.2.1. Reação de fosforilação *in vitro* monitorizada através da espectroscopia de RMN

A fosforilação é uma das modificações pós-traducionais mais frequentes pois representa um importante mecanismo regulador das vias de sinalização das células eucarióticas. É um processo reversível realizado por proteínas cinases. O mecanismo consiste na adição de um grupo fosfato ( $\text{PO}_4$ ) a um grupo polar de um aminoácido da proteína. Esta adição aumenta a polaridade da proteína e, muitas vezes, permite a sua interação com outras moléculas. Um aminoácido fosforilado consegue ligar moléculas capazes de interagir com outras proteínas e, consequentemente, montar ou desmontar complexos proteicos.<sup>45</sup>

Os resíduos mais fosforilados são as serinas, treoninas e tirosinas (figura 1.6). As experiências 2D  $^1\text{H}$ - $^{15}\text{N}$  são particularmente adequadas para identificar fosforilações nestes resíduos, especialmente nos resíduos de serina e treonina, uma vez que fosforilações nestes aminoácidos induzem alterações significativas nos desvios químicos das amidas da cadeia principal.<sup>83</sup> Essas mudanças nos desvios químicos são causadas principalmente pelas ligações por pontes de hidrogénio entre os prótons da amida da cadeia principal e do grupo fosfato que é adicionado no processo de fosforilação.<sup>83-85</sup> A fosforilação em resíduos de tirosina não induz alterações tão grandes nos desvios químicos das amidas da cadeia principal, devido à posição mais distal do grupo hidroxilo, em comparação com o mesmo grupo nos resíduos de serina e treonina (figura 1.6).<sup>83,86</sup>



**Figura 1.6- Fosforilação em resíduos de Serina, Treonina e Tirosina.** A vermelho está representado o grupo fosfato ( $\text{PO}_4$ ) adicionado a cada resíduo pelas proteínas cinases aquando da reação de fosforilação. Figura adaptada de <sup>83</sup>.

Por outro lado, a fosforilação das tirosinas resulta numa substancial alteração nos desvios químicos das ressonâncias dos  $\text{CH}_\epsilon$ , funcionando como indicadores inequívocos da presença de tirosinas fosforiladas. Contudo, o mapeamento dos locais de fosforilação utilizando as ressonâncias  $^1\text{H}$ - $^{13}\text{C}$  das cadeias laterais dos aminoácidos é difícil. Esse mapeamento é normalmente conseguido através das mudanças nos desvios químicos das amidas da cadeia principal dos aminoácidos que se situam em torno dos respetivos resíduos de tirosina

modificados, uma vez que estes, geralmente, exibem maiores diferenças de desvios químicos, que as próprias tirosinas fosforiladas.<sup>83</sup> Assim, a detecção através de RMN da fosforilação das tirosinas e o mapeamento dos locais de fosforilação por combinação das experiências de correlação de 2D  $^1\text{H}$ - $^{13}\text{C}$  e  $^1\text{H}$ - $^{15}\text{N}$  é bastante eficaz.

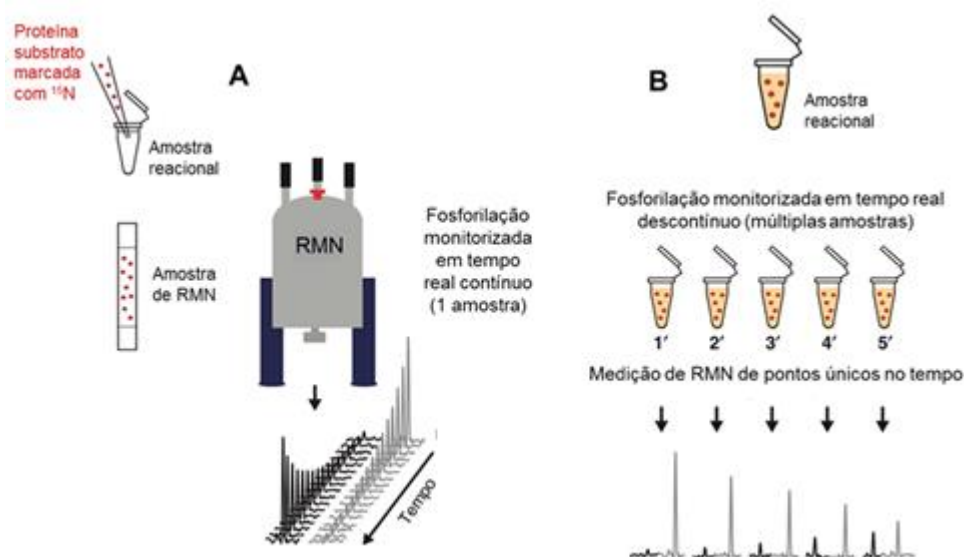
Ao longo deste trabalho, tentamos reconstituir a versão fosforilada do efector essencial à infecção por EPEC e EHEC: Tir. Embora o significado biológico da fosforilação das proteínas seja amplamente apreciado, existe uma carência de métodos para detetar essas modificações *in vitro*. A espectroscopia de RMN foi a técnica utilizada, pois é adequada para detetar a fosforilação de uma proteína *de novo* porque os sinais de ressonância da mesma são altamente sensíveis às mudanças induzidas por essa modificação.<sup>75,83,87,88</sup> Outra vantagem da utilização da espectroscopia de RMN é o estabelecimento progressivo dessas modificações poder ser monitorizado de maneira não invasiva e sem interrupções. Desta forma, as reações de fosforilação nas proteínas podem ser analisadas de maneira contínua e os eventos de modificação gradual podem ser dissecados.<sup>88,89</sup> Estas características juntas, tornam a espectroscopia de RMN uma técnica eficaz para o estudo desta modificação pós-traducional em proteínas.

#### 1.2.2.2. Fosforilação *in vitro* monitorizada em tempo real contínuo e descontínuo

Para monitorizar as reações de fosforilação de proteínas *in vitro*, as proteínas que funcionam como substrato da reação marcadas  $^{15}\text{N}$  são adicionadas à mistura reacional com a cinase não marcada, ATP e  $\text{Mg}^{2+}$ , bem como outros cofatores necessários à atividade da cinase. A fosforilação de locais individuais da proteína substrato leva a alterações nos desvios químicos ao longo do tempo da reação, que podem ser detetadas consecutivamente através de experiências de RMN consecutivas e resolvidas em tempo real (figura 1.7A).<sup>89</sup>

Ao registar essas experiências de RMN resolvidas em tempo real, os eventos de modificação são monitorizados de maneira contínua. Como os sinais de RMN de resíduos da proteína substrato não modificados e modificados são detetados lado a lado, e porque a espectroscopia de RMN é uma técnica inerentemente quantitativa, as mudanças nas intensidades dos sinais fornecem informações específicas locais sobre os níveis de modificação individuais e como estes evoluem ao longo do tempo.<sup>88,89,90</sup> A maior desvantagem da espectroscopia de RMN em estudos de fosfo-deteção de resíduos com a monitorização a ser feita em tempo real de forma contínua é a sensibilidade inerentemente da técnica. Como resultado, estes ensaios requerem concentrações de proteína substrato na gama dos  $\mu\text{M}$ . Nesta concentração, um espectro de RMN 2D requer cerca de 4 minutos para ser concluído. Portanto, os pontos de medição individuais da reação de fosforilação contínua são normalmente definidos com 4 minutos de distância uns dos outros. Isto pode ser um fator limitante quando as reações de fosforilação são rápidas. Existem algumas maneiras de lidar com esta situação: no caso mais simples, a reação de fosforilação é desacelerada usando menos cinase na reação *in vitro*. Mudanças na temperatura da reação e no pH também podem ser equacionadas para diminuir as atividades específicas das cinases.<sup>83,89</sup>

De outra maneira, existe um outro método que também pode ser usado para obter informações qualitativas e quantitativas sobre os estados de fosforilação do substrato em intervalos de tempo mais curtos, onde as reações são mais rápidas. Esse método foi o utilizado neste trabalho, e consiste numa reação de fosforilação *in vitro* monitorizada em tempo real descontínuo. Neste método, a proteína substrato marcada com  $^{15}\text{N}$  é, da mesma forma, adicionada à mistura com a cinase não marcada, o ATP e os cofatores da enzima. Depois são feitas medições de RMN de alíquotas de reações de fosforilação descontínuas, isto é, são retiradas alíquotas da reação ao longo do tempo e as mesmas são colocadas a elevadas temperaturas, interrompendo a reação. A monitorização e quantificação dos locais e níveis de fosforilação é alcançada através da medição de espectros 2D  $^1\text{H}$ - $^{15}\text{N}$  de RMN das diferentes alíquotas, aos diferentes tempos de extensão reacional (figura 1.7B).<sup>89</sup> É importante notar que este procedimento só é adequado para o estudo de IDPs ou pequenos péptidos, pois a inativação a altas temperaturas desnatura e normalmente precipita as proteínas com estrutura tridimensional.

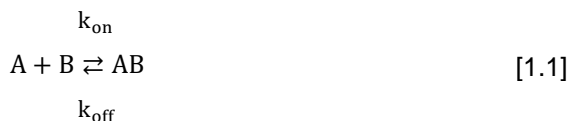


**Figura 1.7- Esquema geral dos procedimentos de RMN para monitorizar reação de fosforilação de proteína *in vitro* em tempo real. (A)- Procedimento para a aquisição de dados de RMN da reação de fosforilação monitorizada em tempo real contínuo (B)- Estratégia para aquisição de dados de RMN da reação de fosforilação monitorizada em tempo real de forma descontínua. Figura adaptada de <sup>89</sup>.**

### 1.2.2.3. Caracterização de interações proteína-proteína através de espectroscopia de RMN

As IDPs, em parte graças à sua flexibilidade estrutural, estão envolvidas em sistemas de regulação, estabelecendo interações com diversas moléculas. Estas interações são na sua maioria transientes, tornando a sua deteção e a sua caracterização um desafio. A espectroscopia de RMN é a técnica por excelência para estudar este tipo de interações, pois é bastante sensível, mesmo a interações fracas. De igual forma, esta técnica permite monitorizar estes fenómenos de associação em solução com resolução atômica, conseguindo retirar informações termodinâmicas e das zonas das proteínas envolvidas nessas interações.

Nesta secção, será tomando como exemplo um sistema simples de interação reversível entre duas proteínas A e B com um local de interação, originando um complexo AB, de acordo com a equação 1.1:



A caracterização destes sistemas de interação proteína-proteína, ocorre geralmente pela titulação de uma das proteínas não marcada (e.g. proteína B) para a solução que contém a outra, marcada isotopicamente com  $^{15}\text{N}$  ou  $^{13}\text{C}$  (e.g. proteína A). A cada adição de titulante (proteína B), são adquiridos espectros de correlações heteronucleares (normalmente  $^1\text{H}$ - $^{15}\text{N}$  HSQC). Desta forma, consegue-se monitorizar a passagem de estado entre a proteína A livre e complexada com a proteína B, formando o complexo AB. É importante que as duas proteínas sejam preparadas no mesmo tampão para minimizar os possíveis artefactos da experiência.

Os desvios químicos são muito sensíveis a mudanças de ambiente químico, o que torna possível a deteção de interações mesmo que essas sejam muito fracas. Tendo em conta o sistema apresentado, os espectros  $^1\text{H}$ - $^{15}\text{N}$  HSQC da proteína titulada marcada isotopicamente (proteína A) contêm as ressonâncias das correlações NH dos seus resíduos. Como a proteína B é titulada na solução da proteína A, as perturbações no espectro da última refletem a mudança de ambiente químico dos grupos NH do complexo formado pela interação destas duas proteínas. Assim sendo, a monitorização dos desvios químicos em cada ponto da titulação permite tirar informação acerca das regiões presentes na interface das interações, pois os desvios químicos desses resíduos serão os mais afetados.<sup>67,91</sup>

Considerando o sistema reversível simples apresentado na equação 1.1, a proteína titulada irá dar origem a duas espécies com diferentes frequências de ressonância ( $\omega_A$  e  $\omega_{AB}$ ) e, consequentemente, a uma diferença de frequências de ressonâncias, definida por  $\Delta\omega = \omega_A - \omega_{AB}$ .<sup>92,93</sup> Contudo, o aparecimento de diferentes estados da proteína no espectro de RMN depende da constante de dissociação ( $K_d$ ) e da constante de permuta de reação ( $k_{ex}$ ), definidas por:

$$K_d = \frac{[A][B]}{[AB]} \quad [1.2]$$

$$k_{ex} = k_{on}[B] + k_{off} \quad [1.3]$$

Para uma interação controlada por difusão (onde o local de interação está estereoquimicamente acessível) os valores de  $k_{on}$  são constantes e da ordem de  $10^9 \text{ M}^{-1}\text{s}^{-1}$ . Existe, portanto, uma relação aproximada entre  $K_d$  e  $k_{on}$ , em que  $k_{off} \approx 10^9 K_d$ .<sup>94</sup>

Estas considerações são importantes para perceber e distinguir os diferentes resultados obtidos nos espectros de RMN nos diferentes sistemas de interação proteína-proteína. No



diagrama esquemático da figura 1.8 estão apresentados alguns desses resultados. De notar que os espectros dessa figura contêm apenas três picos de correlação heteronuclear N-H, representando uma amostra simples de picos, comparativamente ao que seria um espectro real, muito mais complexo e com picos muitas vezes sobrepostos.

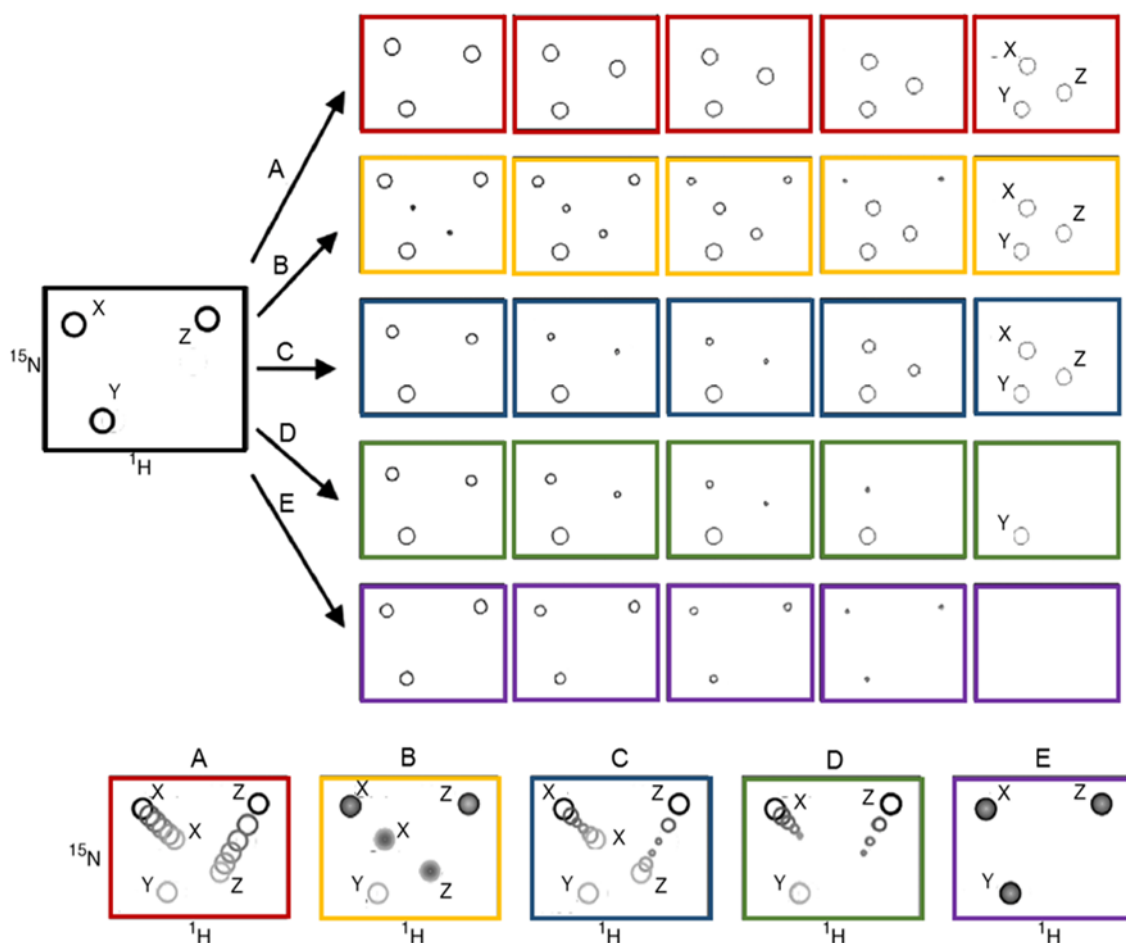
As permutas químicas estão divididas em três categorias: permuta lenta, permuta intermédia e permuta rápida. Esta divisão é baseada na magnitude de  $k_{ex}$ , em relação ao tamanho da diferença das frequências das ressonâncias (ou desvios químicos) dos núcleos da proteína livre A e complexada AB ( $\Delta \omega$ ).<sup>92,93</sup>

Quando  $k_{ex}$  é muito maior do que a diferença dos desvios químicos entre o estado livre e complexado ( $k_{ex} \gg \Delta \omega$ ), a permuta é considerada rápida na escala de tempo de RMN e para um determinado sistema de *spin* apenas se observa uma ressonância que se move suavemente da posição da proteína A no estado livre em direção a uma nova posição no espectro, referente à proteína no estado complexado AB, com pouca ou nenhuma alteração de intensidade de sinal (figura 1.8A).<sup>92-95</sup>

Por outro lado, como se observa na figura 1.8B, quando  $k_{ex}$  é muito menor do que a diferença dos desvios químicos entre estados ( $k_{ex} \ll \Delta \omega$ ), a permuta é lenta na escala de tempo da espectroscopia de RMN e as ressonâncias de cada estado podem ser observadas individualmente. A intensidade dos sinais diminui progressivamente à medida que a proteína A é titulada com quantidades crescentes da proteína B. Consequentemente, aparecem sinais cuja intensidade vai aumentando de forma progressiva, correspondentes ao estado complexado da proteína A. Dessa maneira, a um conjunto de ressonâncias em regime de permuta lenta, podem ser determinados os tamanhos relativos da população dos diferentes estados a partir das intensidades dos sinais, podendo a interação entre as duas proteínas ser seguida diretamente pelas alterações das intensidades do estado complexado (*bound state*).<sup>92-95</sup>

Em ambos os casos de permuta descritos, os gráficos da alteração dos desvios químicos ou da alteração das intensidades (dependendo do regime de permuta) em função do rácio molar entre titulante e titulado ( $[proteína\ B]/[proteína\ A]$ ) produzem hipérboles que podem ser ajustadas a curvas que permitem calcular valores de constantes termodinâmicas, como o  $K_d$ .<sup>94,96</sup>

Para o regime de permuta intermédia ( $k_{ex} = \Delta \omega$ ), também conhecido como coalescência, uma combinação dos dois fenómenos anteriores pode ser observada e podem observar-se alterações progressivas de desvios químicos acompanhadas pelo aumento ou diminuição da intensidade inicial de sinal, com a recuperação da intensidade do pico ao longo da titulação (figura 1.8C). A posição das ressonâncias neste regime é bastante incerta e a interpretação quantitativa é muito complicada, não sendo viável na maioria das vezes.<sup>92-95</sup>



**Figura 1.8- Diagrama esquemático dos espectros 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC de uma titulação de um complexo, da esquerda para a direita, da proteína marcada A com concentração crescente de proteína B (ligando não marcado).** À esquerda está representado o espectro da proteína A nativa, livre de titulante, como três picos das correlações NH da cadeia principal para resíduos fictícios X, Y e Z. Para cada caso, os grupos NH dos resíduos X e Z estão perto da interação entre as proteínas e o resíduo Z está distante da mesma. O aumento da concentração de titulante está representado numa escala de cinza. **(A)-** Titulação em regime de permuta rápida. **(B)-** Titulação em regime de permuta lenta. **(C)-** Titulação em regime de permuta intermédia. **(D)-** Titulação em regime de permuta rápida/intermédia. **(E)-** Titulação com a formação de um complexo invisível através da espectroscopia de RMN. Figura adaptada de <sup>95</sup>.

As representações apresentadas são, de facto, simplistas. Na prática, podemos observar comportamentos de permuta rápida, intermédia e lenta na mesma experiência. O que é comum a todos os regimes de permuta é o alargamento dos sinais do espectro quando há a formação do complexo entre as duas proteínas. Esse alargamento deve-se ao aumento da velocidade do decaimento da magnetização transversal, associado ao maior tempo de correlação rotacional ( $\tau_c$ ) do complexo, em comparação com a proteína livre.<sup>95</sup>

Por vezes, os resultados obtidos são menos diretos. O caso da figura 1.8D é um desses exemplos, onde o comportamento da interação se manifesta de forma semelhante ao do regime de permuta intermédia. Contudo, a progressiva diminuição da intensidade dos picos não se recupera no ponto final da titulação, isto é, no máximo da concentração de titulante (proteína B).

A interpretação destes resultados pode ser feita de duas formas: **i)** o complexo está em regime de permuta intermédia, mas o máximo alcançável de concentração da proteína B é insuficiente para saturar a proteína A. Para resolver este problema, a medição do valor da constante  $K_d$  deve ser obtida por outros métodos, como a Calorimetria de Titulação Isotérmica (ITC, do inglês *Isothermal Titration Calorimetry*); **ii)** o desaparecimento dos picos pode representar o espectro do complexo totalmente formado, havendo o intercâmbio entre duas ou mais conformações do mesmo.<sup>94,95,97</sup> Finalmente, na figura 1.8E está representada uma situação em à medida que a titulação prossegue, a maioria dos picos diminui de intensidade sem nenhuma alteração nos desvios químicos. Tal é explicado pelo complexo AB formado ser “invisível” na experiência de RMN.<sup>95</sup>



## **2. Objetivos**



Tir possui quatro regiões que mimetizam sequências consenso presentes em proteínas do hospedeiro (regiões ITIM). Estas regiões compreendem tirosinas passíveis de sofrer fosforilação. Quando esta modificação pós-traducional ocorre, Tir adquire a capacidade de mediar o recrutamento de proteínas do hospedeiro, nomeadamente proteínas com domínios SH2, de forma a controlar os processos celulares do hospedeiro, pela sua subversão. De entre essas proteínas, destaca-se a fosfatase SHP-1. A interação entre estas duas proteínas é conhecida, contudo, os mecanismos moleculares e a forma como a mesma acontece não o são. Desta forma, a presente dissertação tem como principal objetivo contribuir para a clarificação das dinâmicas moleculares da interação entre a região C-terminal de Tir (C-Tir) e SHP-1.

- Em primeiro lugar, a concretização deste objetivo pressupõe a sobre-expressão e a purificação destas proteínas: C-Tir e os domínios SH2 de SHP-1.

Conforme referido, para que a interação ocorra, é necessário que os resíduos de tirosinas das regiões ITIM estejam fosforilados. Desta maneira, o estudo desta modificação pós-traducional demonstra ser importante para a compreensão da interação entre estas duas proteínas. Assim sendo, um dos objetivos da presente dissertação passa pela reconstituição do padrão de fosforilação de C-Tir por ação de uma cinase de tirosina (Fyn). Para tal, pretendemos:

- Implementar um método de monitorização da fosforilação, através de espectroscopia de RMN;
- Estudar a cinética de fosforilação e a sua extensão, através do mapeamento específico dos aminoácidos fosforilados, através de espectroscopia de RMN.
- Explorar as consequências estruturais da fosforilação a nível local em C-Tir, através da análise de desvios químicos, recorrendo uma vez mais à espectroscopia de RMN.

Uma vez estudada a fosforilação, estaríamos em condições para avançar para o seguinte objetivo: Estudar as bases de reconhecimento molecular da interação entre C-Tir e domínios SH2, utilizando como modelo SHP-1. Para alcançar este objetivo pretendemos realizar ensaios de interação entre as mesmas, através de espectroscopia de RMN, para:

- Quantificar a força da interação relativa entre os diferentes sítios de fosforilação ( $K_d$ );
- Analisar a especificidade da interação.

Planeámos estudar em paralelo a interação entre as duas proteínas no estado não-fosforilado, para confirmar e averiguar se a interação entre as proteínas só acontece quando Tir está no estado fosforilado.





### **3. Materiais e Métodos**



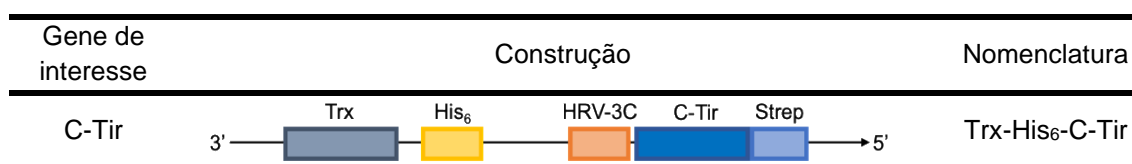
## 3.1. Produção de proteína recombinante

### 3.1.1. Proteína C-Tir

A clonagem, a expressão e purificação da proteína C-Tir foi realizada de acordo com protocolos definidos e otimizados anteriormente no laboratório. (Vieira M. et al, artigo em preparação).

O gene c-tir que codifica os resíduos 388-550 da proteína Tir da estirpe EPEC (código Uniprot B7UM99) foi clonado no vetor de expressão pHTP8 (NZYtech), em trabalho anterior realizado no laboratório. Este vetor permite a expressão da proteína recombinante em fusão com Tio redoxina (Trx) na região N-terminal. Trx é uma proteína termoestável amplamente utilizada para promover aumento de solubilidade.<sup>98</sup> A montante desse promotor de solubilidade Trx, o vetor contém uma cauda de seis resíduos de histidina (His<sub>6</sub>-tag), permitindo a purificação da proteína recombinante por cromatografia de afinidade por iões imobilizados (IMAC, do inglês *Immobilized Metal Ion Affinity Chromatography*) e uma sequência de reconhecimento da protease HRV-3C, possibilitando a posterior separação da proteína de interesse dos parceiros de fusão. A construção utilizada na presente dissertação contém ainda uma Strep-tag, uma cauda de 8 aminoácidos, clonada artificialmente na região C-terminal do gene, por forma a facilitar a purificação da proteína por cromatografia de afinidade de elevada especificidade (tabela 3.1). O mapa do vetor e a sequência da proteína estão apresentados no anexo (figura 7.1 e tabela 7.1, respetivamente).

**Tabela 3.1- Construção do gene C-Tir EPEC no vetor pHTP8.** De notar que os tamanhos dos genes e das tags ilustrados não estão à escala.



#### 3.1.1.1 Expressão da proteína isotopicamente marcada

Células competentes de estirpe BL21 Star (DE3) pLysS de *E. coli* foram transformadas com o plasmídeo descrito acima, adicionando-se 1 µL do plasmídeo a 50 µL de células competentes. A mistura foi incubada em gelo durante 30 minutos, posteriormente incubada a 42 °C durante 40 segundos e, de seguida, colocada novamente em gelo durante 5 minutos. Adicionaram-se 500 µL de meio de cultura LB (NZYtech) à amostra e a mesma foi incubada a 37 °C durante 60 minutos com agitação a 350 rpm (AccuTherm™, Labnet International). O produto obtido foi plaqueado em placas de LB-agar, previamente suplementadas com 50 µg/mL de canamicina (NZYtech) e 34 µg/mL de cloranfenicol (ACROS Organics). As placas foram posteriormente incubadas a 37 °C durante a noite.

Uma colônia isolada foi inoculada em 60 mL de meio de cultura LB, suplementado com os mesmos antibióticos e o mesmo foi incubado a 37 °C, com agitação a 180 rpm durante a noite (Infors HT Minitron).

De seguida, o pré-inóculo foi adicionado ao meio de cultura do crescimento bacteriano, composto por 2 L de meio de cultura Luria Bertani (LB), suplementado com os mesmos antibióticos. Este foi incubado a 37 °C, com agitação a 180 rpm até se atingir um valor de densidade ótica (OD)<sub>600nm</sub> de 1,0. Nessa altura, as células foram centrifugadas durante 15 minutos, a 4 °C e a 8000 rpm (Rotor JA-10, Avanti J-25I, Beckman Coulter) e posteriormente transferidas para 2 L de meio M9 (composição em anexo na tabela 7.12). Para a expressão da proteína recombinante uniformemente marcada com <sup>15</sup>N, o meio de cultura M9 foi suplementado com <sup>15</sup>NH<sub>4</sub>Cl como única fonte de azoto. A expressão com marcação isotópica simultânea com <sup>15</sup>N e <sup>13</sup>C, foi realizada transferindo-se as células para o meio de cultura M9 com a fonte de azoto referida e <sup>13</sup>C-glucose como única fonte de carbono. A expressão de C-Tir foi induzida com 1 mM IPTG (NZYtech) durante 24 horas, a 20 °C e 120 rpm. As células foram recolhidas através de uma centrifugação de 15 minutos, a 4 °C e a 8000 rpm (Rotor JA-10, Avanti J-25I, Beckman Coulter).

#### 3.1.1.2 Purificação da proteína isotopicamente marcada

Os *pellets* provenientes da expressão em *E. coli*, foram inicialmente ressuspensos em tampão de lise (50 mM Tris-HCl pH 7,5, 150 mM NaCl, 5 mM EDTA, 1 mM PMSF, 15 mM β-mercaptoetanol, 100 μL de cocktail inibidor de proteases (PIC), 1 mg DNase e 1.5 mg lisozima) e colocados a 90 °C, durante 30 minutos (AccuTherm™, Labnet International).

De seguida, o lisado foi ultracentrifugado durante 45 minutos, a 4 °C e a 42000 rpm (rotor 70ti, Optima™ LE-80K, Beckman Coulter), sendo recolhida a fração solúvel. A fração insolúvel foi solubilizada num tampão com 50 mM Tris-HCl pH 7,5, 150 mM NaCl, 5 mM EDTA, 1 mM PMSF, 15 mM β-mercaptoetanol, 1% Triton X-100 e 0.5 M Ureia, e foi ultracentrifugada durante 30 minutos, a 4 °C e a 42000 rpm. De seguida, o sobrenadante foi recuperado e o pellet foi novamente ressuspensionado em tampão de lise e ultracentrifugado durante 45 minutos, a 4 °C e a 42000 rpm. O sobrenadante foi uma vez mais recolhido. As frações solúveis recolhidas após cada ultracentrifugação foram analisadas por SDS-PAGE.

Juntaram-se as amostras que continham a proteína de interesse e as mesmas foram purificadas através de uma cromatografia de afinidade, utilizando uma coluna Strep-TactinXT™ (5 mL) (iba), cuja matriz é composta por Streptavidina, que permite uma interação forte e específica com a Strep-tag, presente na região C-terminal da construção. A coluna, previamente equilibrada em tampão de ligação (50 mM Tris-HCl pH 7,5, 150 mM NaCl, 1 mM EDTA), foi adaptada a uma bomba peristáltica (Masterflex I/P® Easy-Load®). O extrato celular foi carregado na coluna e de seguida, procedeu-se ao corte das *tags* Trx e His<sub>6</sub> (presentes na região N-terminal da construção) da proteína de interesse. Esta clivagem enzimática foi realizada diretamente na coluna, por incubação com a protease HRV-3C, durante a noite, a 4 °C. Posteriormente, a coluna foi lavada com 5 volumes de tampão de ligação e após a separação dos constituintes que não




apresentavam afinidade para a matriz da coluna, procedeu-se à eluição, com tampão de eluição (100 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA e 50 mM biotina). As frações recolhidas ao longo deste passo cromatográfico foram analisadas por SDS-PAGE.

Juntaram-se as frações de elevada pureza e integridade num único volume, o qual foi concentrado utilizando um Amicon (Merck, 5 kDa). A concentração final foi determinada através da Lei de Lambert-Beer e da medição da  $A_{280\text{nm}}$ , no espectrofotómetro NanoDrop™ 2000 (Thermo Scientific).

### 3.1.2. Proteína SHP-1

O gene da sequência completa de SHP-1 *wild type* (WT) que codifica os resíduos 1-595 foi adquirido através do repositório Addgene no vetor pGEX-2T (#8594). Os genes das versões truncadas NC-SH2 e C-SH2 que por sua vez codificam os resíduos 1-217 e 104-217 respetivamente, foram adquiridos através do mesmo repositório no vetor pGEX6P1 (#46496 e #46497, respetivamente). Estes vetores permitem a expressão das proteínas de interesse em fusão com Glutathione S-transferase (GST), uma proteína presente na região N-terminal da construção, que promove o aumento da solubilidade e permite a purificação da proteína recombinante por cromatografia de afinidade, funcionando dessa forma como *tag* de solubilidade e afinidade.<sup>99</sup> A montante dessa proteína, estes vetores contêm ainda um local de clivagem para a protease HRV-3C que possibilita a separação da proteína de interesse do parceiro de fusão (tabela 3.2).

**Tabela 3.2- Construções dos genes C-SH2, NC-SH2 e SHP-1 WT nos vetores pGEX6P1 e pGEX-2T.**  
De notar que os tamanhos dos genes e das *tags* não estão à escala.

| Gene de interesse | Construção |  |    | Nomenclatura |
|-------------------|------------|--|----|--------------|
| C-SH2             | 3' —       |  | 5' | GST-C-SH2    |
| NC-SH2            | 3' —       |  | 5' | GST-NC-SH2   |
| SHP-1 WT          | 3' —       |  | 5' | GST-SHP-1    |

#### 3.1.2.1. Clonagem dos genes dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1

Foram idealizadas diferentes construções com o objetivo final de conseguir expressar e purificar os domínios N-SH2, C-SH2 e NC-SH2 de SHP-1. Para tal, os genes que codificam para cada um destes domínios foram clonados em diferentes vetores compostos por diferentes *tags* de solubilidade e afinidade. O gene C-SH2, compreendendo os resíduos 104-217 foi clonado no vetor pHTP8 e os genes N-SH2 e NC-SH2, compreendendo os resíduos 1-107 e 1-217, foram clonados nos vetores pHTP8, pHTP10 (NZYtech) e pCoofy18 (Addgene: # #43975), usando os *kits NZYEasy Cloning & Expression VIII e X* (NZYtech) e o *kit Phusion High-Fidelity PCR* (NEB),

respetivamente. A sequência dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1 estão apresentados no anexo (tabela 7.1).

O vetor pHTP10 tem a mesma constituição que o pHTP8, diferindo apenas na proteína que promove a solubilidade, apresentando NusA em vez de Trx. O vetor pCoofy18 é constituído apenas por uma cauda de dez histidinas (His<sub>10</sub>-tag) na região N-terminal da construção. De notar que em todas as construções desenhadas, foi clonada uma Strep-tag na região C-terminal dos genes. Os mapas dos vetores estão presentes na figura 7.1, em anexo.

O método de clonagem utilizado foi a reação de clonagem independente de sequência e ligação (SLIC, do inglês *Sequence and Ligation Independent Cloning*). Contrariamente às abordagens tradicionais, este método não recorre a enzimas de restrição. Neste método, são adicionadas extensões nas extremidades 5' dos *primers* utilizados para a amplificação dos genes alvo. Essas extensões de cadeia simples, designadas por *non-priming overhangs*, contêm uma sequência de nucleótidos complementar ao vetor onde se pretende clonar o gene de interesse, permitindo a hibridação do gene no vetor aquando da realização da reação de clonagem.

Os genes que codificam para os diferentes domínios foram amplificados, utilizando o plasmídeo GST-NC-SH2 como molde, através de PCR (realizada no termociclador T100™, BIO-RAD). Os programas de amplificação, a composição das misturas reacionais e os *primers* utilizados, encontram-se apresentados em anexo (tabela 7.2, 7.3 e 7.8, respetivamente). A eficiência da amplificação foi avaliada através de um gel de agarose 1% (m/v) (ver composição em anexo na tabela 7.11).

O DNA amplificado foi purificado, com recurso ao kit *NZYGel Pure* (NZYtech). O DNA purificado foi quantificado utilizando o espectrofotómetro NanoDrop™ 2000 (Thermo Scientific). A quantidade ideal de DNA amplificado necessária à reação de clonagem foi calculada para cada construção.

Na clonagem nos vetores pHTP8 e pHTP10, a quantidade ótima de DNA, calculada de acordo com as instruções do kit utilizado, foi misturada com cada um dos vetores linearizados, com a enzima DNA polimerase e com a mistura reacional. Toda a mistura reacional foi colocada no termociclador, estando a composição de cada reação de clonagem e os seus parâmetros presentes na tabela 7.4, em anexo.

Para a clonagem no vetor pcoofy18, o DNA amplificado foi misturado com o vetor linearizado, com a enzima recA (NEB) e com a sua mistura reacional (NEB), sendo a mistura posteriormente incubada a 37 °C, durante 30 minutos. A composição da reação de clonagem encontra-se apresentada na tabela 7.5, em anexo.

As células competentes TOP10 foram transformadas com os produtos da reação de clonagem. A 50 µL de células competentes, adicionaram-se 2 µL do conteúdo de cada produto da reação e incubou-se em gelo durante 30 minutos. Passado esse tempo, seguiu-se uma incubação a 42 °C (AccuTherm™, Labnet International) durante 60 segundos e novamente em gelo durante 2 minutos. De seguida, adicionaram-se 250 µL de meio SOC (NZYtech), previamente aquecido, e incubou-se a 37 °C durante 60 minutos com agitação a 350 rpm (AccuTherm™, Labnet International). Passado esse tempo, as células transformadas foram

plaqueadas em LB-agar, previamente suplementadas com 50 µg/mL de canamicina. As placas foram incubadas a 37 °C durante a noite.

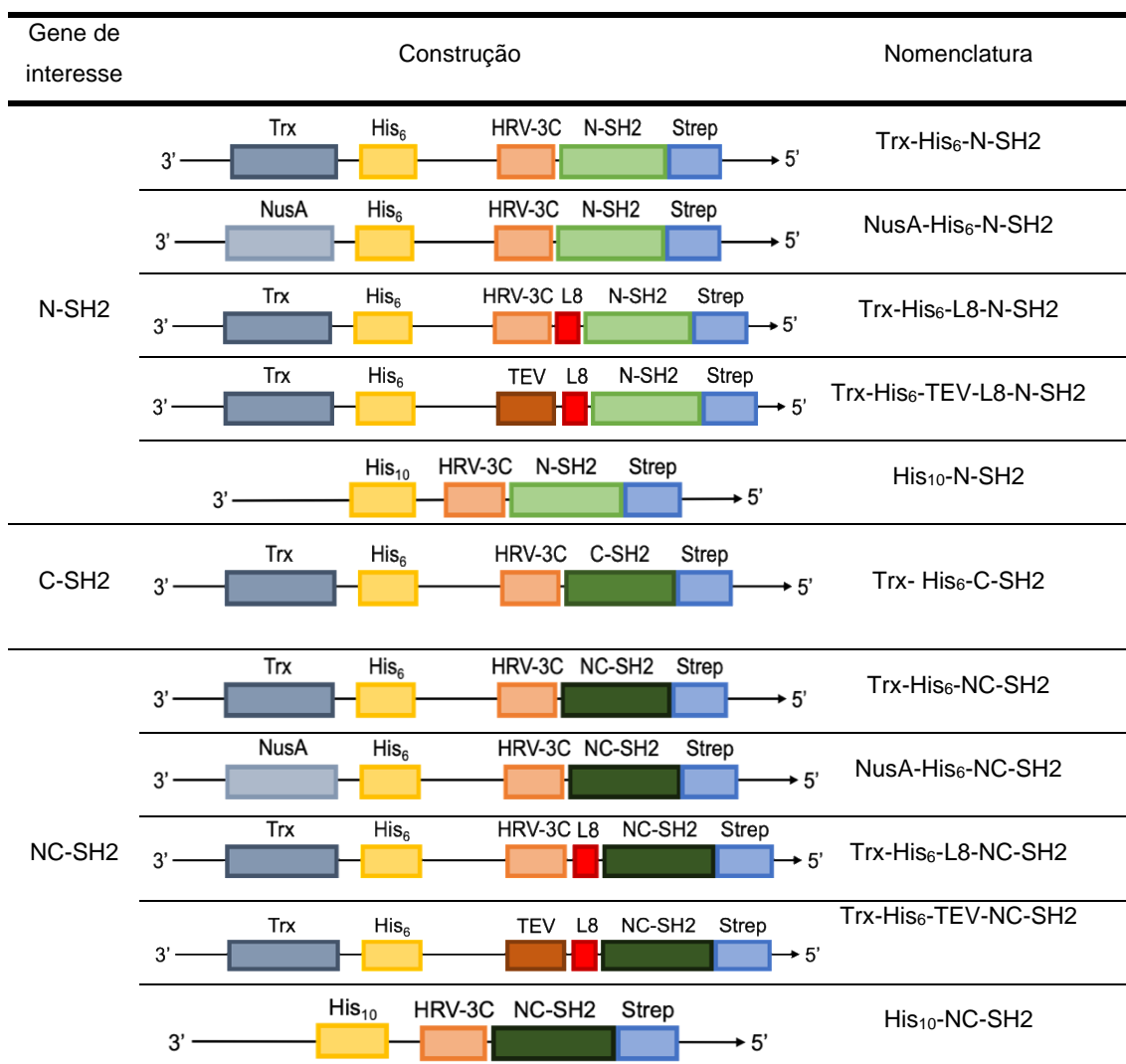
Para identificar as colónias positivas de cada construção foi realizado um *screening* das colónias recombinantes. Cada colónia foi inoculada em 11 mL de meio de cultura LB, suplementado com 50 µg/mL de canamicina, com agitação a 180 rpm durante a noite (Infors HT Minitron). De cada inóculo, utilizou-se 1 mL para efetuar o *screening*, sendo aquecido a 90 °C durante 10 minutos (AccuTherm™, Labnet International). O *screening* foi realizado através de PCR, em que o DNA fervido foi utilizado como molde na mistura reacional. Os *primers* usados foram os mesmos que já haviam sido utilizados na amplificação de cada gene de interesse. Os programas de amplificação e a composição de cada mistura reacional encontram-se presentes na tabela 7.6 e 7.7, respetivamente, em anexo. O *screening* das colónias recombinantes foi analisado através de um gel de agarose 1%.

Depois de verificadas quais as colónias recombinantes positivas, o restante volume de inóculo (10 mL) foi utilizado para extrair e isolar o DNA, usando o kit *NZYMiniPrep* (NZYtech). O DNA foi quantificado utilizando o espectrofotómetro NanoDrop™ 2000 (Thermo Scientific) e os plasmídeos foram confirmados através de sequenciação.

Na tabela 3.3, encontram-se representadas as diferentes construções utilizadas na presente dissertação para a produção dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1.

**Tabela 3.3- Construções realizadas com os genes de interesse: N-SH2, C-SH2 e NC-SH2 de SHP-1.**

De notar que os tamanhos dos genes e das *tags* não estão à escala e que o L8 é a representação de um *linker* de 8 resíduos, clonado artificialmente entre a região de clivagem pela protease e os genes de interesse.



### 3.1.2.2. Testes de expressão dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1

Foram realizados testes de expressão para as diferentes construções com os genes codificantes para os domínios N-SH2, C-SH2 e NC-SH2 de SHP-1. Para tal, foram testadas diferentes condições com o intuito de otimizar e aumentar o rendimento da expressão dos genes desses domínios na fração solúvel. As condições testadas incluíram: diferentes estirpes de *E. coli* como células hospedeiras, meios de cultura, OD de indução, temperaturas, tempos de indução e diferentes concentrações de indutor (IPTG). Na tabela 3.4, estão apresentadas todas as condições testadas.



Todas as estirpes de células competentes testadas foram transformadas com os diferentes plasmídeos, seguindo o protocolo descrito em 3.1.1.1. As células e os plasmídeos foram selecionados na presença dos respectivos antibióticos (apresentados na tabela 3.4).

**Tabela 3.4- Condições testadas para a sobre-expressão das diferentes proteínas de interesse: domínio N-SH2, C-SH2 e NC-SH2 de SHP-1.**

| Plasmídeos                  |  |  |
|-----------------------------|--|--|
| N-SH2                       | Trx-His <sub>6</sub> -N-SH2 (pHTP8)          | Ampicilina (100 µg/mL)   |
|                             | Trx-His <sub>6</sub> -L8-N-SH2 (pHTP8)       | Canamicina (50 µg/mL)  |
|                             | NusA-His <sub>6</sub> -N-SH2 (pHTP10)        | Canamicina (50 µg/mL)  |
|                             | Trx-His <sub>6</sub> -TEV-L8-N-SH2 (pHTP10)  | Canamicina (50 µg/mL)  |
|                             | His <sub>10</sub> -N-SH2 (pCoofy18)          | Canamicina (50 µg/mL)  |
| C-SH2                       | Trx-His <sub>6</sub> -3C-C-SH2 (pHTP8)       | Canamicina (50 µg/mL)  |
| NC-SH2                      | Trx-His <sub>6</sub> -NC-SH2 (pHTP8)         | Canamicina (50 µg/mL)  |
|                             | Trx-His <sub>6</sub> -L8-NC-SH2 (pHTP8)      | Canamicina (50 µg/mL)  |
|                             | NusA-His <sub>6</sub> -NC-SH2 (pHTP10)       | Canamicina (50 µg/mL)  |
|                             | Trx-His <sub>6</sub> -TEV-L8-NC-SH2 (pHTP10) | Canamicina (50 µg/mL)  |
|                             | His <sub>10</sub> -NC-SH2                    | Canamicina (50 µg/mL)  |
| Condições testadas          |  |  |
| Células hospedeiras         |  |  |
|                             | BL21 (DE3)                                   | —  |
|                             | BL21 Star (DE3)pLysS                         | Cloranfenicol (34 µg/mL)   |
|                             | Tuner (DE3)                                  | —  |
|                             | ArticExpress (DE3)                           | Gentamicina (20 µg/mL)   |
|                             | Rosetta (DE3)                                | Cloranfenicol (34 µg/mL)   |
|                             | BL21-Codon Plus (DE3)                        | Cloranfenicol (34 µg/mL)   |
|                             | C41 (DE3)                                    | Cloranfenicol (34 µg/mL)   |
| [IPTG] (mM)                 |  | 0; 0,05; 0,1; 0,25; 0,5; 1,0   |
| OD <sub>indução</sub>       |  | 0,4-0,5  |
|                             |  | 0,6-0,8  |
|                             |  | 0,8-1,0  |
|                             |  | 1,0-1,2  |
| Temperatura de indução (°C) |  | 11, 12, 18, 20, 25, 37   |
| Tempo de indução (h)        |  | 2, 4, 6, 12, 16, 20, 24  |
| Meio de cultura             |  | LB, M9, TB   |
| Método de ruptura celular   |  | <i>BugBuster®</i> 1x (lise química)<br><i>French-press</i> (lise mecânica) |

O volume dos crescimentos e das alíquotas retiradas dependeu do método como ocorreria a lise celular. No caso dos testes em que a lise celular foi feita recorrendo à solução *BugBuster®*

1x, realizaram-se crescimentos de 25 mL de meio de cultura e recolheram-se alíquotas de 1 mL aos diferentes tempos de indução. Estas foram centrifugadas a 12000 xg, durante 10 minutos (LLG-uniCFUGE 5) e os *pellets* foram ressuspensos em 150 µL da solução *BugBuster*® 1x (50 mM Tris-HCl pH 7,5, 150 mM NaCl, 1 mM PMSF, 1 mM DTT e solução *BugBuster*® 1x (Novagen)). Com o intuito de separar as frações solúveis e insolúveis realizou-se uma nova centrifugação, a 12000 xg durante 10 minutos. Por outro lado, quando a lise foi conseguida através de *french-press*, o volume dos crescimentos foi de 500 mL de meio de cultura e foram recolhidas alíquotas de 25 mL. As amostras foram centrifugadas a 4000 rpm, durante 20 minutos e os *pellets* foram ressuspensos em 25 mM Tris-HCl pH 7,5, 150 mM NaCl, 5 mM EDTA, 10 mM β-mercaptoetanol, 25 µL cocktail inibidor de proteases, 1 mg DNase e 1.5 mg lisozima. A esse passo seguiu-se uma ultracentrifugação durante 1 hora a 4 °C e 42000 rpm (rotor TLA 100.3, Optima™TLX, Beckman Coulter), separando as frações solúveis e insolúveis.

Em ambos os casos, as frações solúveis e insolúveis foram analisadas por SDS-PAGE. De forma a avaliar qualitativamente a sobre-expressão das diferentes proteínas de interesse nas diferentes condições testadas, os volumes das amostras colocadas nos poços foram normalizados pelo valor de OD<sub>600nm</sub>.

### 3.1.2.3. Expressão dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1

Os testes de expressão permitiram otimizar as condições de expressão dos domínios de interesse.

Deste modo, a expressão do domínio C-SH2 de SHP-1 foi conseguida através da transformação das células competentes de estirpe BL21 Star (DE3)pLysS com o plasmídeo Trx-His<sub>6</sub>-C-SH2, seguindo o protocolo descrito em 3.1.1. Tanto a transformação como o crescimento foram realizados na presença de 50 µg/mL de canamicina e 34 µg/mL de cloranfenicol, devido à resistência do plasmídeo e das células, respetivamente. O crescimento realizado foi de 2 L de meio de cultura LB, a 37 °C e com agitação a 180 rpm (Infors HT Minitron). A expressão da proteína foi induzida após se ter atingido um valor de OD<sub>600nm</sub> de 0,6 pela adição de 0,1 mM IPTG e de seguida as células foram incubadas a 20 °C com agitação a 120 rpm. Passadas 24 h do momento de indução, as células foram recolhidas através de uma centrifugação de 15 minutos, a 4 °C e a 8000 rpm (rotor JA-10, Avanti 25I, Beckman Coulter).

A expressão dos domínios N-SH2 e NC-SH2 ocorreu de forma semelhante. Desta feita, as células competentes de estirpe BL21-Codon Plus (DE3) foram transformadas com os plasmídeos desses genes nos vetores pHTP8 e pHTP10. O crescimento foi realizado da mesma maneira, sendo a expressão das proteínas de interesse induzida do mesmo modo com 0,1 mM IPTG, mas apenas quando se atingiu uma OD<sub>600nm</sub> de 0,8. As células foram incubadas a 18 °C com agitação a 120 rpm e após 16 horas do momento da indução as células foram igualmente recolhidas.

#### 3.1.2.4. Purificação do domínio C-SH2 de SHP-1

As células foram inicialmente ressuspensas em tampão de lise (composição referida em 3.1.2.2.) e realizou-se a lise celular recorrendo a *french-press*. O extrato celular foi ultracentrifugado durante 1 hora e 30 minutos, a 4 °C e a 42000 rpm (rotor 70ti, Optima™ LE-80K, Beckman Coulter), separando a fração solúvel da insolúvel.

O primeiro passo de purificação consistiu numa cromatografia de afinidade. A fração solúvel foi carregada na coluna Strep-TactinXT™ (5 mL) (iba) previamente equilibrada em tampão de ligação (50 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA e 1 mM DTT). A coluna estava adaptada a uma bomba peristáltica (Masterflex I/P® Easy-Load®). De seguida, a coluna foi lavada com 5 volumes de tampão de ligação e a eluição foi conseguida através da passagem de 4 volumes de tampão de eluição (100 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA, 1 mM DTT e 50 mM biotina). As frações recolhidas foram analisadas por SDS-PAGE.

Juntaram-se as frações eluídas que continham a proteína de interesse, para clivagem enzimática com a protease HRV-3C, com o intuito de remover as *tags* Trx e His<sub>6</sub> da mesma. Para tal, o tampão foi trocado para 50 mM Tris-HCl pH 7,5, 150 mM NaCl, 1 mM EDTA e 1 mM DTT, utilizando um Amicon (Merck, de 10 kDa) e foram adicionados 250 µL da protease HRV-3C (1 mg/mL). A digestão ocorreu durante a noite, a 4°C e a eficiência da clivagem foi analisada por SDS-PAGE.

A etapa seguinte foi a realização de uma IMAC. Para tal, adaptou-se uma coluna de níquel (Ni<sup>2+</sup>) HisTrap™ (5 mL) (GE Healthcare), previamente equilibrada com tampão de ligação (50 mM Tris-HCl pH 7,5, 150 mM NaCl, 1 mM DTT e 20 mM imidazol), a uma bomba peristáltica (Masterflex I/P® Easy-Load®). Os iões Ni<sup>2+</sup> imobilizados na matriz desta coluna interagem com a cauda de histidinas da His<sub>6</sub>-*tag*, permitindo a separação das *tags* Trx e His<sub>6</sub> (presentes na região N-terminal da construção) da proteína de interesse. A amostra de proteína foi carregada na coluna e a mesma foi lavada com 5 volumes de tampão de ligação, permitindo a eluição da proteína de interesse. De seguida, as *tags* Trx e His<sub>6</sub> foram eluídas com 5 volumes de tampão de eluição (50 mM Tris-HCl pH 7,5, 150 mM NaCl, 1 mM DTT e 0.5 M imidazol). As frações recolhidas foram analisadas por SDS-PAGE e as que continham a proteína de interesse juntaram-se. A amostra foi concentrada, utilizando um Amicon Vivaspin 20 (GE Healthcare, de 5 kDa) para um volume final de 500 µL.

A etapa final da purificação consistiu numa cromatografia de exclusão molecular, um método cromatográfico que separa as moléculas em solução pelo seu tamanho molecular. A amostra de proteína foi injetada numa coluna Superdex 200 Increase 10/300 GL (24 mL) (GE Healthcare), cujos limites de exclusão são de 10-600 kDa, acoplada a um sistema AKTA™ Purifier 10 (GE Healthcare), sendo a A280nm seguida através do software UNICORN Start (GE Healthcare). A coluna foi equilibrada com 2 volumes de 20 mM HEPES pH 6,8, 150 mM NaCl, 1 mM DTT e a amostra injetada, sendo a eluição isocrática e a absorvância monitorizada a 280 nm. As frações recolhidas foram analisadas por SDS-PAGE e juntaram-se as frações que continham a proteína de interesse. A amostra foi concentrada utilizando um Amicon Vivaspin 20 (GE Healthcare, de 5

kDa). A concentração final foi determinada através da Lei de Lambert-Beer e da medição da  $A_{280\text{nm}}$ , no espectrofotômetro NanoDrop™ 2000 (Thermo Scientific).

### 3.1.2.5. Tentativas de purificação dos domínios N-SH2 e NC-SH2 de SHP-1

A lise celular e a consequente ultracentrifugação dos domínios N-SH2 e NC-SH2 de SHP-1 ocorreram da forma descrita anteriormente para o domínio C-SH2. Foram realizadas diversas tentativas de purificação destes, nas diferentes construções disponíveis. As diferentes abordagens para purificar os domínios encontram-se apresentadas na tabela 3.5. Os tampões utilizados nas presentes tentativas de purificação, foram os mesmos que já haviam sido utilizados antes na purificação do domínio C-SH2 (consultar secção 3.1.2.4.).

**Tabela 3.5- Condições testadas nas diversas tentativas de purificação dos domínios N-SH2 e NC-SH2 de SHP-1.** A **negrito** estão representados os domínios de interesse, em cada construção.

| Plasmídeo                                   | Coluna                 | Clivagem enzimática<br>(protease, local de corte) |
|---|------------------------|---|
| Trx-His <sub>6</sub> - <b>N-SH2</b>         | Strep-TactinXT™ (5 mL) | HRV-3C, dentro e fora da coluna                   |
| Trx-His <sub>6</sub> - <b>NC-SH2</b>        | HisTrap™ (5 mL)        | HRV-3C, dentro da coluna                          |
| NusA-His <sub>6</sub> - <b>N-SH2</b>        | Strep-TactinXT™ (5 mL) | HRV-3C, fora da coluna                            |
| NusA-His <sub>6</sub> - <b>NC-SH2</b>       |                        |   |
| Trx-His <sub>6</sub> -L8- <b>N-SH2</b>      | Strep-TactinXT™ (5 mL) |   |
| Trx-His <sub>6</sub> -L8- <b>NC-SH2</b>     |                        |   |
| Trx-His <sub>6</sub> -TEV-L8- <b>N-SH2</b>  | Strep-TactinXT™ (5 mL) | TEV, fora da coluna                               |
| Trx-His <sub>6</sub> -TEV-L8- <b>NC-SH2</b> |                        |   |

## 3.2. Caracterização de C-Tir através da espectroscopia de RMN

Todas as experiências de espectroscopia de RMN foram adquiridas num instrumento de alto campo de 18.8T Bruker Avance II+ a operar a uma frequência de próton de 800 MHz (CERMAX), equipado com uma sonda criogénica TXI de tripla ressonância. As amostras foram medidas em tubos de RMN de 3 mm, com 10% D<sub>2</sub>O e 20 µL DSS (Sigma), este último para servir como calibração dos desvios químicos <sup>1</sup>H. Os dados foram todos adquiridos a 10 °C, processados com o *software* Bruker TopSpin 3.6.2™ (Bruker BioSpin) e posteriormente analisados com o *software* CARA.<sup>100</sup>

### 3.2.1. Atribuição das ressonâncias de C-Tir fosforilada

Segundo o descrito na literatura, a fosforilação dos resíduos de tirosina presentes nas regiões ITIM é um requisito para a interação entre C-Tir e os domínios SH2. Dessa maneira, o estudo desta modificação pós-traducional e dos seus efeitos a nível local em C-Tir é fundamental para o entendimento de todo o processo de interação com as proteínas do hospedeiro.

A atribuição das ressonâncias da cadeia principal de aminoácidos da proteína C-Tir não-fosforilada foi realizada anteriormente no laboratório (*Vieira et al, artigo em preparação*). De forma a atribuir as ressonâncias da cadeia principal da proteína fosforilada de forma inequívoca, foi necessário recorrer a experiências multidimensionais (2D,3D) de correlação entre  $^1\text{H}$ ,  $^{13}\text{C}$  e  $^{15}\text{N}$ .

#### 3.2.1.1. Preparação da amostra e aquisição dos dados

O tampão da proteína duplamente marcada com  $^{15}\text{N}$  e  $^{13}\text{C}$  foi trocado para 20 mM HEPES pH 7.5, 150 mM NaCl, 12 mM  $\text{MgCl}_2$  e 5.5 mM ATP, utilizando um Amicon (Merck, 5 kDa). A reação de fosforilação das tirosinas pela adição da cinase Fyn (Thermo Fisher Scientific) (1,93  $\mu\text{M}$ ) à amostra de proteína (1.1 mM), durante a noite, a 25 °C. A quantidade de cinase Fyn utilizada teve em conta a atividade específica da mesma a 25 °C e foi adicionada num excesso de 5 vezes, para garantir a fosforilação completa dos quatro locais de fosforilação. Uma vez realizada a reação, o tampão da proteína foi trocado para 20 mM HEPES pH 6.8, 150 mM NaCl, 12 mM  $\text{MgCl}_2$  e 5.5 mM ATP, utilizando um Amicon (Merck, 5 kDa).

Como controlo da reação, foi adquirido um espectro  $^1\text{H}$ - $^{15}\text{N}$  HSQC da proteína não-fosforilada (969  $\mu\text{M}$ ). De seguida, foram adquiridas as experiências 2D HSQC e as experiências 3D HNCO e HNCACB da proteína fosforilada. Os parâmetros de cada experiência encontram-se apresentados na tabela 3.6. O sucesso da reação de fosforilação foi verificado através da observação das alterações nos desvios químicos dos quatro resíduos de tirosinas e dos seus resíduos vizinhos.

**Tabela 3.6- Parâmetros das experiências de RMN.** NS indica o número de scans, TD o número de incrementos, JW a janela espectral e d1 o intervalo de tempo entre scans.

| Experiência | Sequencia de pulso    | NS | Núcleos         | TD   | JW (Hz)  | JW (ppm) | Centro (Hz) | Centro (ppm) | d1 (s) |
|-------------|-----------------------|----|-----------------|------|----------|----------|-------------|--------------|--------|
| HNCO        | hncogp3d (Bruker)     | 16 | $^1\text{H}$    | 2048 | 8802,81  | 10,998   | 3753,55     | 4,690        | 1      |
|             |                       |    | $^{15}\text{N}$ | 64   | 1885,72  | 23,250   | 9609,96     | 118,5        |        |
|             |                       |    | $^{13}\text{C}$ | 64   | 2314,70  | 11,500   | 35279,92    | 175,3        |        |
| HNCACB      | hncacbgp3d (Bruker)   | 72 | $^1\text{H}$    | 2048 | 8802,81  | 10,998   | 3753,55     | 4,690        |        |
|             |                       |    | $^{15}\text{N}$ | 56   | 1703,23  | 21,000   | 9528,87     | 117,5        |        |
|             |                       |    | $^{13}\text{C}$ | 96   | 11068,82 | 55,000   | 8150,35     | 40,50        |        |
| HSQC        | hsqcetf3gpsi (Bruker) | 10 | $^1\text{H}$    | 2048 | 11160,71 | 13,945   | 3753,55     | 4,690        |        |
|             |                       |    | $^{15}\text{N}$ | 256  | 1906,0   | 23,500   | 9650,51     | 119,0        |        |

### 3.2.1.2. Processamento dos dados

A atribuição das ressonâncias da proteína fosforilada beneficiou do facto da atribuição das ressonâncias da proteína não-fosforilada ter sido realizada anteriormente (*Vieira M. et al, artigo em preparação*). Assim, a reatribuição sequencial das ressonâncias de C-Tir fosforilada foi realizada recorrendo à experiência 3D HNCACB, cujo espectro contém correlações inter e intraresiduais, pois correlaciona o NH de um resíduo (*i*) com os C $\alpha$  e C $\beta$  do próprio resíduo e do resíduo anterior (*i-1*). Para fazer a atribuição do carbono carbonilo (CO), foi utilizada a experiência 3D HNCO, uma vez que esta correlaciona o HN de um resíduo (*i*) com o CO do resíduo anterior (*i-1*). Para mais detalhe sobre as experiências 3D adquiridas consultar secção 1.2.2.

As consequências estruturais da fosforilação na proteína foram analisadas através do cálculo dos desvios químicos secundários ( $\Delta\delta$ ). Esses valores representam as diferenças dos valores dos desvios químicos observados em relação a valores dos mesmos em conformação *random coil* e dão informação acerca das propensões de elementos de estrutura secundária. Foram utilizados valores desvios químicos em conformação *random coil* corrigidos e ajustados para IDPs.<sup>101</sup> A fim de corroborar e aumentar a robustez das predições, para além do cálculo dos  $\Delta\delta^{C\alpha}$  e  $\Delta\delta^{C\beta}$  e da diferença entre eles, recorreu-se ao método *Neighbour Corrected Structural Propensity Calculator* (ncSPC) que estima preferências estruturais utilizando de forma combinada os desvios químicos de  $^{13}C^O$ ,  $^{13}C^\alpha$ ,  $^{13}C^\beta$ ,  $^{15}N$  e  $^1H$  atribuídos a cada aminoácido da cadeia principal da proteína fosforilada.<sup>82</sup> Desta previsão, os aminoácidos com valores entre -0,1 e 0,1 foram considerados adotarem conformações *random coil* e valores superiores a 0,1 e inferiores a -0,1 foram considerados como propensões a hélices- $\alpha$  e folhas- $\beta$ , respetivamente (consultar secção 1.2.2., para mais detalhe acerca da forma como são calculados os  $\Delta\delta$  de cada átomo).

### 3.2.2. Reação de fosforilação de C-Tir *in vitro* monitorizada em tempo real descontínuo

Com o intuito de estudar a cinética e a extensão da reação de fosforilação das tirosinas das regiões ITIM mimetizadas por C-Tir, foi realizada uma reação de fosforilação *in vitro* monitorizada em tempo real descontínuo, que consiste em retirar diferentes alíquotas a diferentes tempos de reação, após se ter adicionado a cinase à proteína isotopicamente marcada com  $^{15}N$ . A monitorização e quantificação dos locais de fosforilação individuais foi alcançada através da medição de espectros de RMN aos diferentes tempos de reação.

#### 3.2.2.1. Preparação da amostra e aquisição dos dados

O tampão da proteína isopticamente marcada com  $^{15}N$  foi trocado para 20 mM HEPES pH 7.5, 150 mM NaCl, 12 mM MgCl<sub>2</sub> e 2 mM ATP, utilizando um Amicon (Merck, 5 kDa). A cinase Fyn (1,01  $\mu$ M) foi adicionada à amostra de proteína (543,17  $\mu$ M). A reação ocorreu a 25 °C, à

qual foram retiradas alíquotas (50  $\mu$ M) a diferentes tempos de reação: 2,4, 8, 16, 24, 32, 48 e 64 minutos. Como controlo desta experiência, utilizamos uma amostra de proteína não-fosforilada à mesma concentração das restantes (50  $\mu$ M). A reação foi inativada a 95°C (AccuTherm™, Labnet International), durante 5 minutos, sendo as alíquotas posteriormente congeladas em azoto líquido e armazenadas a -80°C até serem adquiridos os dados de RMN. Antes da aquisição das experiências, o tampão da proteína foi trocado para 20 mM HEPES pH 6.8, 150 mM NaCl, 12 mM MgCl<sub>2</sub> e 5.5 mM ATP. As experiências <sup>1</sup>H-<sup>15</sup>N HSQC aos diferentes tempos de reação foram adquiridas utilizando os mesmos parâmetros de aquisição da amostra referência, os quais se encontram apresentados anteriormente na tabela 3.6.

### 3.2.2.2. Processamento dos dados

Para extrair informações quantitativas sobre os níveis de fosforilação de cada tirosina em específico, foram calculadas as intensidades relativas dos sinais desses resíduos nos espectros de RMN com a razão entre intensidade absoluta desses sinais ( $I_{desf}$ ) e a intensidade dos mesmos no espectro de referência da proteína não fosforilada ( $I_{desf(t=0)}$ ), de acordo com a equação 3.1. Desta forma, conseguimos monitorizar a fração de proteína ainda não-fosforilada ao longo do tempo  $Q(t)$ .

$$Q(t) = \frac{F_{escala} \times I_{desf}(t)}{I_{desf}(t = 0)} \quad [3.1]$$

De notar que o  $F_{escala}$  é o fator de correção que reporta para as diferenças nas condições experimentais entre as amostras aos vários tempos do ensaio e a amostra referência. Para calcular o  $F_{escala}$  foram selecionados sinais de RMN de resíduos que não são afetados pela reação de fosforilação e que por essa razão mantêm a intensidade dos seus sinais ao longo da mesma. De seguida, foram calculadas as razões entre a intensidade de cada um destes sinais ao longo da reação e a intensidade dos mesmos na amostra controlo (não fosforilada). Desta maneira, obteve-se um valor médio de  $F_{escala}$  para cada espectro <sup>1</sup>H-<sup>15</sup>N HSQC adquirido.

Posto isto, a percentagem de fosforilação de cada tirosina ao longo do tempo do ensaio foi calculada seguindo a equação 3.2:

$$\%fosforilação(t) = 1 - Q(t) = 1 - \frac{F_{escala} \times I_{desf}(t)}{I_{desf}(t = 0)} \quad [3.2]$$

As informações quantitativas obtidas através da % de fosforilação de cada tirosina ao longo da reação foram ajustadas a curvas de acumulação específicas desses locais, de forma a poder tirar conclusões dos comportamentos de fosforilação dos resíduos individualmente. O ajuste destas curvas de fosforilação foi conseguido utilizando a equação 3.3, onde os parâmetros  $K_M$  e  $V_{max}$  foram ajustados.

$$\frac{[pS](t)}{[S]_{(t=0)}} = 1 - \left( \frac{K_M}{[S]_{t=0}} \times \ln \left( 1 + \left( e^{\frac{[S]_{t=0}}{K_M}} - 1 \right) \times \exp \left( -\frac{V_{\max}}{K_M} \times t \right) \right) \right) \quad [3.3]$$

A equação 3.3 corresponde a uma aproximação da equação de Michaelis-Menten.<sup>89</sup> Para obter as propriedades e constantes cinéticas exatas da reação de fosforilação, deveriam ser realizadas medições das taxas de modificação inicial a diferentes concentrações de substrato. Face ao ensaio realizado, estes parâmetros ajustados dão informação apenas biológica e semi-quantitativa sobre as diferentes características dos eventos de fosforilação nos diferentes resíduos, que deve ser analisada em termos relativos.

### 3.2.3. Interação entre C-Tir e o domínio C-SH2 de SHP-1

A literatura faz referência à necessidade da fosforilação dos resíduos de tirosina de C Tir para que a sua interação com os domínios SH2 seja possível. Todavia, para averiguar se C-Tir apresenta alguma propensão para interagir com estes domínios sem a necessidade de ser fosforilada nos seus resíduos de tirosina, decidimos estudar a possível interação entre C-Tir não-fosforilada com o domínio C-SH2 de SHP-1.

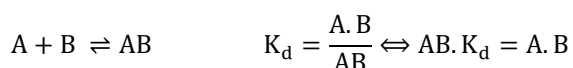
#### 3.2.3.1. Preparação da amostra e aquisição dos dados

O tampão das proteínas foi trocado para 20 mM HEPES pH 6.8, 150 mM NaCl e 1 mM DTT, utilizando um Amicon (Merck, 5 kDa). C-Tir isotopicamente marcada com <sup>15</sup>N (70 μM) foi titulada com 1; 2; 4; 6 e 8 equivalentes do domínio C-SH2 de SHP-1. Foram adquiridas experiências 2D <sup>1</sup>H-<sup>15</sup>N HSQC para cada amostra representativa de cada ponto da titulação realizada. Os parâmetros da experiência encontram-se apresentados anteriormente na tabela 3.6.

#### 3.2.3.2. Processamento dos dados

Graças à atribuição das ressonâncias da cadeia principal de aminoácidos da proteína C-Tir não-fosforilada realizada anteriormente no laboratório foi possível mapear e monitorizar a perturbação das intensidades dos resíduos afetados pela interação referida. A curva de titulação (a diminuição da intensidade relativa de sinal em função do rácio molar [C-SH2]/[C-Tir]) foi ajustada usando a minimização dos mínimos quadrados, num modelo de ligação 1:1, de forma a calcular a constante de dissociação ( $K_d$ ).

Assumindo um modelo simples de estequiometria de ligação 1:1





Durante a titulação, sabemos as concentrações totais de  $A$  e  $B$ , que podemos escrever como  $A_0$  e  $B_0$ . Estas são a soma das suas formas livres e complexadas, isto é:

$$A_0 = A + AB \quad \text{e} \quad B_0 = B + AB$$

Atendendo a que:  $AB = f_b \times A_0$ ;

$$A = (1 - f_b) \cdot A_0$$

$$B = (B_0 - f_b \cdot A_0)$$

em que  $f_b$  representa a fração de proteína complexada

Dessa forma temos:

$$\begin{aligned} (f_b \cdot A_0) \cdot K_d &= [(1 - f_b) \cdot A_0] \times (B_0 - f_b \cdot A_0) \Leftrightarrow \\ \Leftrightarrow f_b \cdot K_d - (B_0 - f_b \cdot A_0 - B_0 \cdot f_b + f_b^2 \cdot A_0) &= 0 \Leftrightarrow \\ \Leftrightarrow f_b &= \frac{(K_d + B_0 + A_0) - \sqrt{(K_d + B_0 + A_0)^2 - 4 \cdot B_0 \cdot A_0}}{2 \cdot A_0} \end{aligned}$$

Sabendo que:  $f_b = \frac{I_{f_{\max}} - I_{\text{obs}}}{I_{f_{\max}} - I_{b_{\min}}}$

a equação que permite calcular  $K_d$  a partir dos valores de intensidade relativa de sinal ( $I_{\text{obs}}$ ), assumindo o modelo descrito é a seguinte:

$$I_{\text{obs}} = I_{f_{\max}} - (I_{f_{\max}} - I_{b_{\min}}) \cdot \frac{(K_d + B_0 + A_0) - \sqrt{(K_d + B_0 + A_0)^2 - 4 \cdot B_0 \cdot A_0}}{2 \cdot A_0} \quad [3.4]$$

em que  $I_{f_{\max}}$  representa o valor de intensidade relativa máximo da proteína no estado livre e  $I_{b_{\min}}$  o valor de intensidade relativa mínimo da proteína complexada quando se atinge a saturação. Os parâmetros ajustáveis são  $K_d$  e  $I_{b_{\min}}$ . De notar que as intensidades relativas ( $I_{\text{obs}}$ ) foram calculadas a partir das intensidades dos sinais retiradas diretamente dos espectros de RMN ao longo da titulação ( $I_{t_{\text{obs}}}$ ), através da equação 3.5, usando como referência as intensidades do espectro da proteína livre ( $I_{f_{\text{obs}}}$ ). O  $F_{\text{escala}}$  foi calculado da mesma forma descrita anteriormente na secção 3.2.2.2.

$$I_{\text{obs}} = \frac{F_{\text{escala}} \times I_{t_{\text{obs}}}}{I_{f_{\text{obs}}}} \quad [3.5]$$

### 3.2.4. Interação entre C-Tir fosforilada e o domínio C-SH2 de SHP-1

Uma vez estudada a fosforilação de C-Tir e feita a atribuição das ressonâncias da cadeia principal da proteína nesse estado, estávamos em condições para avançar com o estudo da interação entre C-Tir fosforilada e o domínio C-SH2 de SHP-1.

#### 3.2.4.1. Preparação da amostra e aquisição dos dados

O tampão de C-Tir marcada isotopicamente com  $^{15}\text{N}$  foi trocado para 20 mM HEPES pH 7.5, 150 mM NaCl, 12 mM  $\text{MgCl}_2$  e 5.5 mM ATP, utilizando um Amicon (Merck, 5 kDa). A reação de fosforilação de C-Tir foi realizada da mesma forma descrita anteriormente em **3.2.1.1.** (1,85  $\mu\text{M}$  de Fyn para 991,1  $\mu\text{M}$  de proteína). De seguida, o tampão de C-Tir e do domínio C-SH2 foi trocado para 20 mM HEPES pH 6.8, 150 mM NaCl, 12 mM  $\text{MgCl}_2$ , 5.5 mM ATP e 1 mM DTT. Como controlo da reação, foi adquirido um espectro  $^1\text{H}$ - $^{15}\text{N}$  HSQC. Uma vez confirmada a fosforilação das suas tirosinas, C-Tir (95  $\mu\text{M}$ ) foi titulada com 0,25; 0,5; 0,75; 1; 2; 4 e 6 equivalentes do domínio C-SH2 de SHP-1. Os espectros  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC foram adquiridos utilizando os mesmos parâmetros apresentados na tabela 3.6.

#### 3.2.4.2. Processamento dos dados

A atribuição das ressonâncias da cadeia principal de C-Tir fosforilada permitiu o mapeamento e a monitorização da perturbação das intensidades dos sinais dos resíduos afetados pela interação referida. As perturbações das intensidades relativas desses sinais foram colocados em função do rácio molar  $[\text{C-SH2}]/[\text{C-Tir}]$  e a curva de titulação foi ajustada usando a minimização dos mínimos quadrados no modelo de interação 1:1 descrito anteriormente, usando também as mesmas equações descritas.

## **4. Resultados e Discussão**



## 4.1. Produção de proteína recombinante

### 4.1.1. Proteína C-Tir

A clonagem do gene C-Tir no vetor de expressão pHTP8 já tinha sido realizada anteriormente no laboratório (*Vieira M. et al, artigo em preparação*). A escolha deste vetor prende-se com o facto deste permitir a expressão da proteína recombinante em fusão com Trx na região N-terminal, uma proteína que promove a solubilidade. A montante da Trx, o vetor contém uma cauda de seis histidinas (*His<sub>6</sub>-tag*), e de seguida apresenta também uma sequência de reconhecimento da protease HRV-3C, importante para a separação de C-Tir dos seus parceiros de fusão. Na região C-terminal, foi também clonada uma cauda de 8 resíduos (*Strep-tag*) de forma a possibilitar a purificação por cromatografia de afinidade de elevada especificidade e eficiência na eliminação de contaminantes (composição da construção presente na tabela 3.1, da secção 3.1.1.).

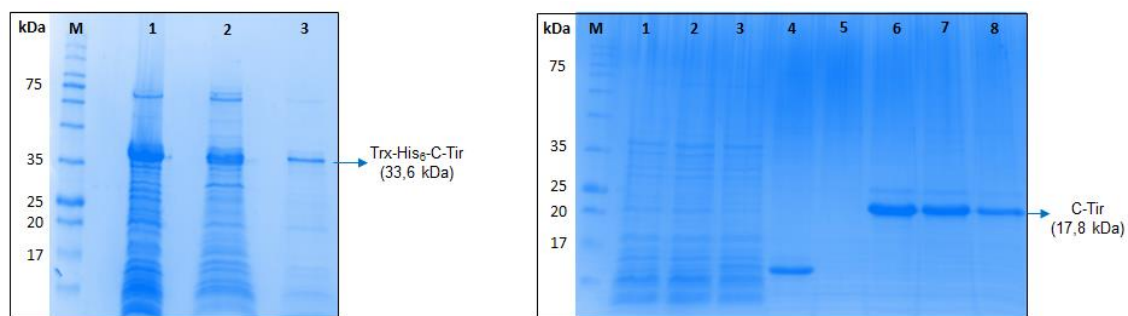
#### 4.1.1.1. Expressão e purificação da proteína isotopicamente marcada

A expressão e purificação da proteína C-Tir isotopicamente marcada foi realizada de acordo com protocolos definidos e otimizados anteriormente no laboratório (*Vieira M. et al, artigo em preparação*).

Após a expressão da proteína recombinante isotopicamente marcada procedeu-se à purificação da mesma. A lise celular a altas temperaturas, designada por *Boiling Lysis*, foi o método utilizado para a ruptura das células. Este método é muitas vezes utilizado para IDPs e é particularmente importante na purificação de C-Tir, pois permite a eliminação de alguns contaminantes e, como C-Tir é uma proteína desordenada e é bastante suscetível à degradação por proteólise, permite a eliminação de proteases presentes no extrato celular.<sup>102</sup> Sendo uma IDP, C-Tir não é gravemente afetada pela desnaturação por temperatura. Desta feita, este método permite a extração e subsequente purificação sem a necessidade de várias etapas de cromatografia ou o uso de inibidores de protease. Após a realização deste processo e a subsequente ultracentrifugação, foi verificado que a proteína recombinante, se encontrava presente em ambas as frações, solúvel e insolúvel. Com o objetivo de obter a maior quantidade possível de proteína, procedeu-se à sua recuperação em corpos de inclusão através de sucessivas lavagens e ultracentrifugações (composição dos tampões e protocolo em 3.1.1.2.). O gel da figura 4.1A, onde é possível observar as frações solúveis resultantes de cada lavagem, demonstra a importância deste processo, uma vez que o mesmo indica a presença de proteína recuperada da fração insolúvel, após cada lavagem.

A purificação da proteína recombinante foi alcançada com recurso a uma única cromatografia de afinidade, utilizando uma coluna Strep-TactinXT™ (5 mL). As frações contendo a proteína foram juntas e injetadas na coluna. Como referido anteriormente, a construção continha uma

Strep-tag na região C-terminal, permitindo uma interação específica com a matriz da coluna, composta por Streptavidina.<sup>103</sup> O corte das tags Trx-His<sub>6</sub> (presentes na região N-terminal da construção) realizou-se diretamente na coluna, pela adição de uma protease (HRV-3C), responsável pelo reconhecimento de uma região específica de aminoácidos, localizada entre as tags Trx-His<sub>6</sub> e a sequência da proteína de interesse. Esta clivagem enzimática ocorreu durante a noite, a 4 °C. De seguida, a coluna foi lavada com tampão de ligação, permitindo a eluição das tags e a eluição da proteína de interesse, foi conseguida pela adição de biotina ao tampão, uma vez que esta compete com a Strep-tag pela ligação à matriz da coluna.



**Figura 4.1- SDS-PAGE (12%) da purificação de C-Tir. (A)-** SDS-PAGE (12%) com as frações solúveis provenientes das sucessivas lavagens e ultracentrifugações. M- Marcador de pesos moleculares NZYColour Protein Marker II; 1- Fração solúvel da 1ª lavagem; 2- Fração solúvel da 2ª lavagem; 3- Fração solúvel da 3ª lavagem. **(B)-** SDS-PAGE (12%) com as frações provenientes do passo de purificação por cromatografia de afinidade. M- Marcador de pesos moleculares NZYColour Protein Marker II; 1 a 3- Frações Flow-through; 4- Fração de lavagem; 5 a 8- Frações de eluição.

As frações recolhidas foram analisadas por SDS-PAGE e encontram-se apresentadas na figura 4.1B. Para a realização dos ensaios pretendidos, foi de extrema importância a obtenção de proteína recombinante pura. Pela análise dos poços 6, 7 e 8 referentes às frações eluídas, da figura 4.1B, é possível verificar a presença de uma banda única, correspondente ao peso molecular da C-Tir (17.8 kDa), indicando o elevado grau de pureza da amostra.

#### 4.1.2. Proteína SHP-1

A presente dissertação visa o estudo da interação entre a região C-terminal de Tir e os domínios de SHP-1 (consultar figura 1.3 da secção 1.1.3.), nomeadamente os domínios SH2 (N-SH2, C-SH2 e NC-SH2). Na fase inicial do trabalho, os genes que codificam esses domínios isolados estavam clonados nos vetores pGEX-2T e pGEX6P1 (composição das construções apresentadas na tabela 3.2 da secção 3.1.2.). Foram realizados testes de expressão dessas construções que indicavam que estes estavam a ser expressos, mas apresentavam-se em corpos de inclusão. Dessa forma, os genes foram clonados em diferentes vetores de expressão, com o intuito de otimizar a produção de proteína recombinante na fração solúvel.

##### 4.1.2.1 Clonagem dos genes dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1

Os genes codificantes para os domínios N-SH2 (378 pb) e NC-SH2 (726 pb) de SHP-1 foram clonados em três vetores de expressão diferentes: pHTP8, pHTP10 e pCoofy18, usando o vetor pGEX6P1 como molde. O gene que codifica o domínio C-SH2 (414 pb) foi clonado apenas no vetor de expressão pHTP8, usando o mesmo molde utilizado na clonagem dos outros domínios.

Estes domínios fazem parte da constituição de uma fosfatase humana (SHP-1) e a expressão dos mesmos durante a dissertação foi realizada num sistema bacteriano. Podem ocorrer algumas dificuldades em expressar genes humanos em organismos bacterianos, pois existem diferenças entre os genes dos mesmos, como os códons usados e os sinais de início de tradução, que ao serem distintos, podem interferir na tradução eficiente do mRNA humano em ribossomas bacterianos.<sup>104</sup>

Muitas vezes, nestas condições, as proteínas são expressas em corpos de inclusão (fração insolúvel). Conceber um protocolo adequado de solubilização e de *refolding* é um desafio árduo e sem garantia de sucesso, uma vez que não há forma de averiguar se o enrolamento da proteína aconteceu corretamente. A estratégia mais comum utilizada para evitar grande parte destes problemas é associar o gene de interesse a um segundo gene bem estabelecido e que se sabe ser bem expresso em sistemas bacterianos, gerando uma proteína de fusão.<sup>105</sup>

Conseguir expressar esses domínios num sistema bacteriano e na fração solúvel foi então uma prioridade neste trabalho. Os vetores de expressão referidos anteriormente foram escolhidos para a clonagem dos genes de interesse pois contêm proteínas utilizadas para promover a solubilidade da proteína recombinante, funcionando como *tag* de solubilidade. O vetor pHTP8 contém na região N-terminal, Tiorredoxina (Trx), uma proteína com 11,7 kDa de *E. coli* conhecida por aumentar significativamente a solubilidade de proteínas pouco solúveis, tornando-as mais estáveis em solução.<sup>93</sup> Da mesma maneira, o vetor pHTP10, contém na região N-terminal, NusA, uma proteína com 54,8 kDa de *E. coli*, utilizada para promover a solubilidade e a estabilidade da proteína recombinante.<sup>106</sup> Tal como o vetor anterior, também este contém uma cauda de seis resíduos de histidinas.

A escolha do vetor de expressão pCoofy18, deveu-se ao facto de este conter apenas uma cauda de dez resíduos de histidinas (His<sub>10</sub>-*tag*), na região N-terminal. Essa cauda apresenta baixo peso molecular e tem um efeito mínimo no enrolamento da proteína de interesse, permitindo trabalhar com a mesma, sem necessidade de remover a cauda de histidinas.

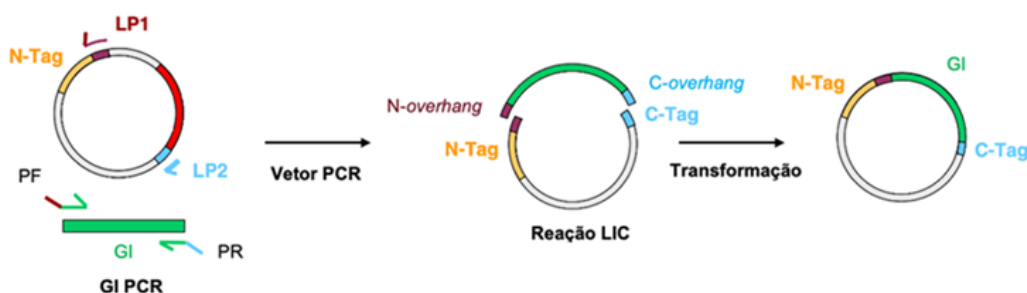
Os métodos tradicionais de clonagem usam enzimas de restrição e enzimas ligase de DNA *in vitro*. A digestão com endonucleases de restrição e ligação aumentam a complexidade dos processos de clonagem de genes exigindo, por exemplo, a seleção de locais de restrição apropriados e etapas de ligação muitas vezes ineficientes. Consequentemente, vários métodos de clonagem independentes de ligação (LIC) foram desenvolvidos por serem mais rápidos, simples e eficientes. Essas estratégias passam por gerar fragmentos de DNA com extremidades complementares em cadeia simples, de forma a permitir a clonagem direcional de qualquer fragmento, sem que sejam necessárias enzimas de restrição ou ligação.

A técnica utilizada neste trabalho foi a clonagem independente de sequência e ligação (SLIC). Segundo a mesma, as extremidades 3' dos primers da PCR são específicas para o DNA molde,

enquanto que as extremidades 5' incorporam caudas específicas complementares ao vetor de expressão onde se pretende clonar o gene de interesse. Ao amplificar esse gene, através da PCR, o seu produto é misturado com o vetor linearizado pretendido e as extremidades de cadeia simples dos primers, as *non-priming overhangs*, vão unir-se eficientemente por complementaridade, após a transformação em *E. coli*, pela recombinação do DNA bacteriano e pela maquinaria de reparo (figura 4.2)<sup>107–110</sup>

Na extremidade 5' dos *primers forward*, podem ser também adicionados alguns nucleótidos para além das extremidades complementares ao vetor, para que os mesmos sejam acrescentados na nova construção. Inicialmente, na clonagem dos genes nos vetores de expressão pHTP8 e pHTP10, foi adicionada uma sequência de nucleótidos, que quando traduzida era reconhecida pela protease HRV-3C. Face ao insucesso da purificação dos domínios N-SH2 e NC-SH2 nessas construções (como se encontrará descrito com maior detalhe na secção 4.1.2.4.), foram desenhadas novas construções para esses domínios, no vetor pHTP8, substituindo essa sequência de nucleótidos, por uma que quando traduzida, passaria a ser reconhecida pela protease TEV. A inclusão dessas sequências reconhecidas por proteases nas diferentes construções seria importante para que, através de uma clivagem enzimática específica, se conseguissem remover as *tags* das proteínas de interesse. Tendo em conta os mesmos problemas de purificação, foram idealizadas outras construções para esses dois domínios no vetor pHTP8. Essas incluíam a inserção de um *linker* de 24 nucleótidos (8 aminoácidos – GGS SGGGS) entre a região de nucleótidos codificantes para aminoácidos reconhecidos pela protease HRV-3C e os nucleótidos dos genes de interesse. A sequência de nucleótidos do *linker* foi adicionada artificialmente ao *primer forward*.

Em todas as clonagens, na extremidade 3' dos *primers reverse*, foi adicionada também uma sequência de nucleótidos antes do codão *stop*, que quando traduzida formaria uma Strep-tag na região C-terminal de cada construção, tal como já tinha sido realizado anteriormente no laboratório, para a construção de C-Tir no vetor de expressão pHTP8.

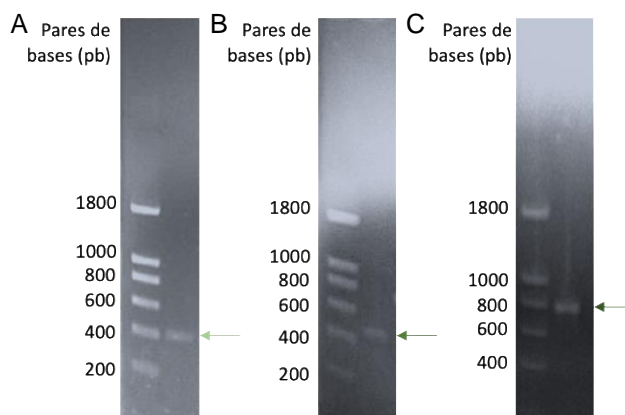


**Figura 4.2- Esquema ilustrativo dos princípios da clonagem por reação LIC.** O vetor é linearizado por PCR com os primer forward LP1 e primer reverse LP2. O gene de interesse (GI) é amplificado por PCR com primers compostos por sequências específicas do gene 5' e 3' mais as extensões (*overhangs*) complementares ao vetor onde se pretende clonar os GI.

Os resultados obtidos nas clonagens dos genes de interesse nos vetores pHTP8, pHTP10 e pCoofy18 foram os mesmos, estando apenas representados de seguida, os que foram obtidos aquando da clonagem dos mesmos no vetor pHTP8.



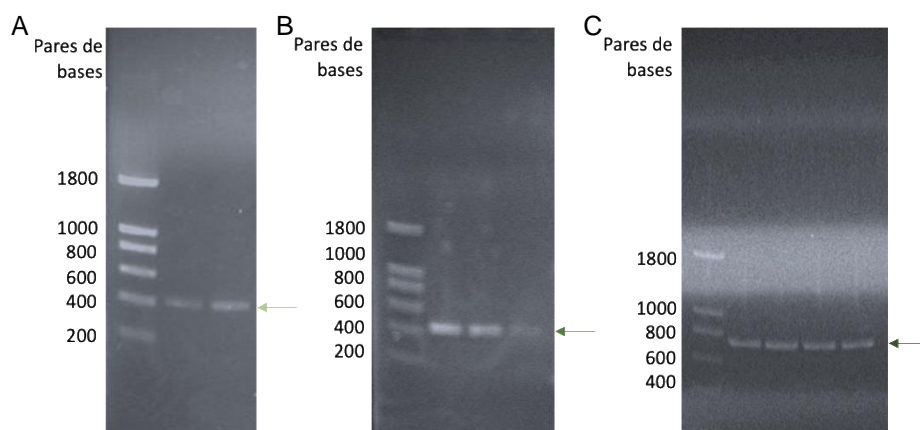
Tal como referido anteriormente, a primeira etapa das clonagens consistiu na amplificação dos genes de interesse através de uma PCR. A eficiência das reações foi avaliada através de géis de agarose 1% (m/v), apresentados na figura 4.3.



**Figura 4.3- Géis de agarose 1% (m/v) da amplificação dos genes de interesse dos domínios de SHP-1. (A)-** Amplificação do gene N-SH2 (378 pb, indicado com a seta verde claro). **(B)-** Amplificação do gene C SH2 (414 pb, indicado com a seta verde). **(C)-** Amplificação do gene NC-SH2 (726 pb, indicado com a seta verde escuro).

Pela análise dos géis, presentes na figura 4.3, é possível observar a existência de uma banda única em cada poço, correspondente ao número de pares de bases (pb) de cada gene amplificado (378 pb, 414 pb e 726 pb para N-SH2, C-SH2 e NC-SH2, respetivamente), verificando desta forma o sucesso das amplificações.

O gene de cada domínio amplificado foi purificado e quantificado no NanoDrop, de forma a calcular e garantir a quantidade ideal para se proceder à reação de clonagem. Aquando da realização da reação de clonagem, o gene amplificado por PCR foi misturado com os vetores linearizados, com a DNA polimerase e com a mistura reacional. O vetor circular recombinante contendo o fragmento de interesse foi obtido transformando o DNA resultante da reação de clonagem em células competentes TOP10. De forma a garantir a inserção dos genes de interesse nos vetores pretendidos, procedeu-se ao *screening* de colónias positivas. Para tal, foram selecionadas colónias isoladas de cada construção e as mesmas foram crescidas num volume pequeno de meio de cultura LB, servindo de molde de DNA na PCR realizada. Nessa reação foram utilizados os mesmos *primers* que já haviam sido utilizados anteriormente na amplificação dos genes. O resultado foi avaliado através de géis de agarose 1% (m/v), apresentados na figura 4.4.



**Figura 4.4- Géis de agarose 1% (m/v) resultantes do screening das colónias recombinantes obtidas depois da reação de clonagem dos genes dos domínios de SHP-1. (A)- Screening das colónias do gene N-SH2 (378 pb, indicado com a seta verde claro). (B)- Screening das colónias do gene C-SH2 (414 pb, indicado com a seta verde). (C)- Screening das colónias do gene NC-SH2 (726 pb, indicado com a seta verde escuro).**

Através da análise dos géis da figura 4.4, é possível verificar que as colónias escolhidas são todas positivas, pois apenas se observa a presença de uma banda em cada poço, cada uma correspondente ao número de pares de bases do respetivo gene de interesse (378 pb, 414 pb e 726 pb, para o gene N-SH2, C-SH2 e NC-SH2, respetivamente). A presença destas bandas indica que houve amplificação dos genes de interesse, o que, por sua vez, indica que esses genes se encontram inseridos nos vetores pretendidos. Os plasmídeos foram confirmados através de sequenciação.

#### 4.1.2.2 Testes de expressão dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1

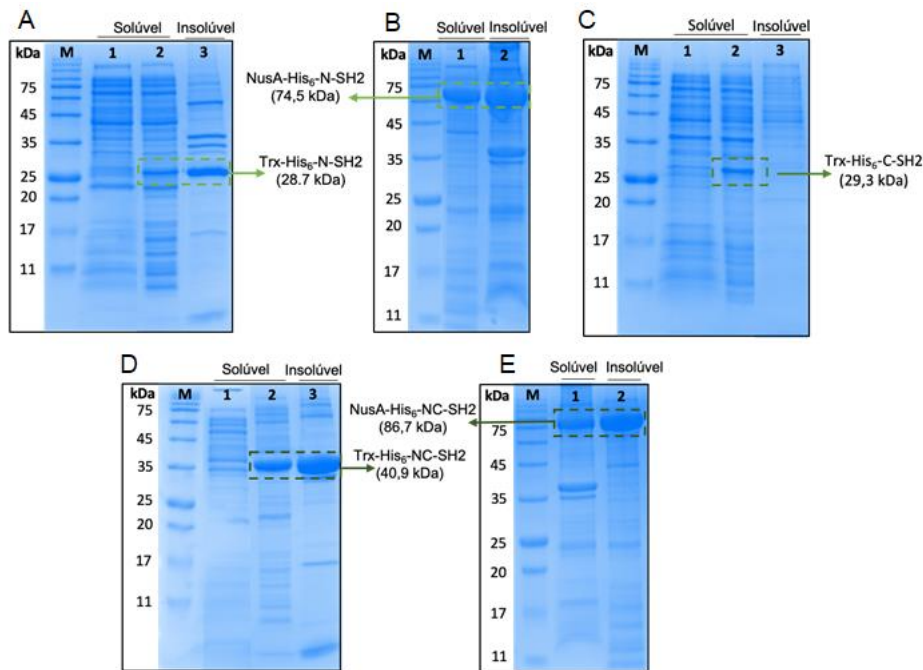
Uma vez realizada a clonagem dos genes codificantes para os domínios de interesse nos vetores de expressão pretendidos, foi necessário proceder a testes de expressão, para otimizar a expressão destes domínios.

A presença de DNA plasmídico recombinante e a expressão de uma proteína recombinante geralmente impõe um gasto de recursos energéticos às células. Essa carga metabólica adicionada reduz frequentemente as taxas de crescimento celular, causa instabilidade e alterações metabólicas e fisiológicas na célula hospedeira, afetando substancialmente os rendimentos funcionais da proteína recombinante alvo.<sup>111,112</sup> A fim de obter o melhor rendimento da expressão das diferentes construções dos domínios de SHP-1, foram testadas diferentes condições, como a temperatura, a composição do meio de cultura, a concentração de indutor, o ponto ideal para induzir a expressão, a sua duração e o próprio hospedeiro bacteriano onde a mesma ocorrerá.

Uma vez feita a análise dos rendimentos obtidos face às condições testadas, as melhores condições para expressar cada um dos domínios de interesse na forma solúvel encontram-se sumariadas na tabela 4.1. Os géis obtidos nessas condições encontram-se apresentados na figura 4.5.

**Tabela 4.1- Condições otimizadas para a expressão dos domínios N-SH2, C-SH2 e NC-SH2 de SHP-1 na fração solúvel.**

| Domínio de interesse | Vetores de expressão | Células hospedeiras   | Meio de crescimento | [IPTG] (mM) | OD <sub>indução</sub> | T <sub>indução</sub> (°C) | t <sub>indução</sub> (horas) |
|----------------------|----------------------|-----------------------|---------------------|-------------|-----------------------|---------------------------|------------------------------|
| N-SH2                | pHTP8/10             | BL21-Codon Plus (DE3) | LB                  | 0.1         | 0.8                   | 18                        | 16                           |
| C-SH2                | pHTP8                | BL21 Star (DE3)pLysS  | LB                  | 0.1         | 0.6                   | 20                        | 24                           |
| NC-SH2               | pHTP8/10             | BL21-Codon Plus (DE3) | LB                  | 0.1         | 0.8                   | 18                        | 16                           |



**Figura 4.5- SDS-PAGE (12%) dos testes de expressão dos domínios N-SH2 nos vetores pHTP8 (A) e pHTP10 (B), C-SH2 no vetor pHTP8 (C) e NC-SH2 nos vetores pHTP8 (D) e pHTP10 (E). (A), (C), (D).** M- Marcador de pesos moleculares NZYColour Protein Marker II; 1- Fração solúvel antes da indução (0h); 2- Fração solúvel depois do tempo de indução; 3- Fração insolúvel depois do tempo de indução. **(B), (E)-** M- Marcador de pesos moleculares NZYColour Protein Marker II; 1- Fração solúvel depois do tempo de indução; 2- Fração insolúvel depois do tempo de indução.

Pela análise dos géis é possível verificar a sobre-expressão de cada um dos domínios de interesse nas condições apresentadas. Enquanto o domínio C-SH2 é expresso unicamente na fração solúvel, os domínios N-SH2 e NC-SH2 apesar de uma parte ser expressa na fração

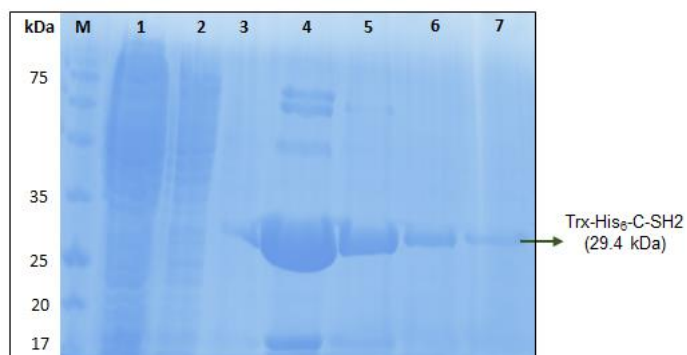
solúvel, a maioria é expressa na fração insolúvel, sob a forma de corpos de inclusão. Apesar de parte dos últimos domínios estarem nestes corpos de inclusão, a quantidade de proteína solúvel seria suficiente para avançar com a purificação dos mesmos.

A expressão dos domínios N-SH2 e NC-SH2 no vetor de expressão pCoofy18 ocorreu sempre na fração insolúvel nas condições testadas. Dessa forma, essas construções foram abandonadas.

#### 4.1.2.3 Expressão e purificação do domínio C-SH2 de SHP-1

Após a expressão do domínio C-SH2 de SHP-1 de acordo com as condições otimizadas, procedeu-se à purificação do mesmo, através de dois passos consecutivos de cromatografia de afinidade, seguidos de uma cromatografia de exclusão molecular.

Depois da lise celular e a subsequente ultracentrifugação terem sido realizadas, a proteína recombinante, presente na fração solúvel, foi purificada através de uma cromatografia de afinidade, utilizando uma coluna Strep-TactinXT™. Tal só é possível, devido a, como já foi referido, esta construção conter uma Strep-tag na região C-terminal, que permite a interação forte e específica com a matriz da coluna. A eluição da proteína foi conseguida através da adição de biotina ao tampão, pois esta molécula compete com a Strep-tag pela ligação à matriz da coluna, libertando a proteína de interesse.



**Figura 4.6- SDS-PAGE (12%) com as frações provenientes do primeiro passo de purificação do domínio C SH2 de SHP-1 por cromatografia de afinidade.** M- Marcador de pesos moleculares NZYColour Protein Marker II; 1- Fração Flow-through; 2- Fração de lavagem; 3 a 7- Frações de eluição.

Este primeiro passo de purificação foi analisado através de um gel SDS-PAGE, apresentado na figura 4.6. É possível confirmar, nos poços 4, 5, 6 e 7, referentes à eluição, a presença da proteína de interesse, com as tags Trx e His<sub>6</sub> na extremidade N-terminal (29,4 kDa). No entanto, apesar da proteína de fusão constituir a banda de maior intensidade, é possível verificar a existência de outras bandas correspondentes a outras proteínas de peso molecular variável, indicando que a proteína de interesse não se encontrava pura.

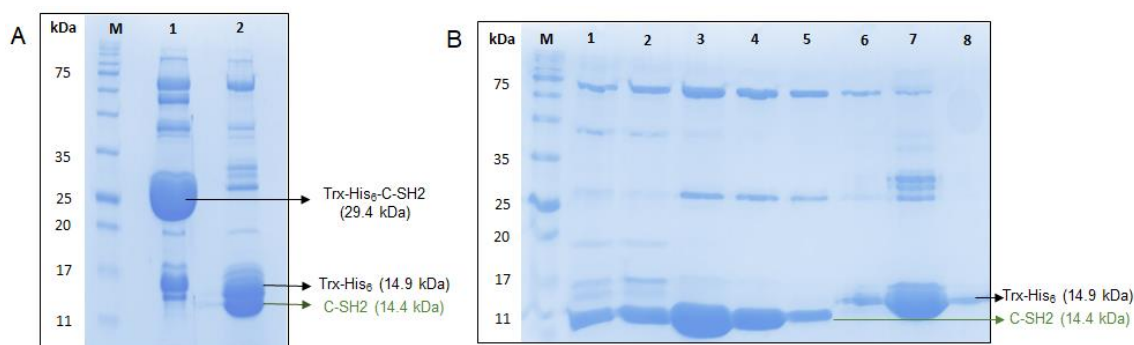
Seguidamente a esta etapa de purificação, que removeu a maioria dos contaminantes, as tags Trx e His<sub>6</sub>, foram cortadas da proteína de interesse através de uma digestão com a protease

HRV-3C. Desta vez, a clivagem enzimática aconteceu fora da coluna, durante a noite, a 4 °C, em 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA e 1 mM DTT.

A digestão com HRV-3C foi analisada através de SDS-PAGE (figura 4.7A) e a eficiência do processo foi comprovada. Conforme é visível no gel, passamos de uma banda correspondente ao peso molecular da proteína fundida com as *tags* (29,4 kDa), para duas bandas, depois da digestão com a protease, muito semelhantes em tamanho, uma com 14,9 kDa e outra com 14,4 kDa, correspondentes às *tags* Trx-His<sub>6</sub> e à proteína de interesse, respectivamente.

Foi realizada uma segunda cromatografia de afinidade para separar as *tags* Trx-His<sub>6</sub> da proteína de interesse. Desta feita, procedeu-se a uma IMAC. Este passo cromatográfico baseia-se no reconhecimento molecular, envolvendo interações específicas e reversíveis, de um ligando alvo a uma molécula imobilizada na coluna. Neste caso, o ligando são as caudas de histidinas, que reconhecem e apresentam elevada afinidade aos íons de níquel na forma de ácido nitriloacético de níquel (Ni-NTA), imobilizados na coluna HisTrap™ High Performance.<sup>113</sup>

As frações recolhidas ao longo desta etapa de purificação foram também analisadas por SDS-PAGE, estando o gel apresentado na figura 4.7B.



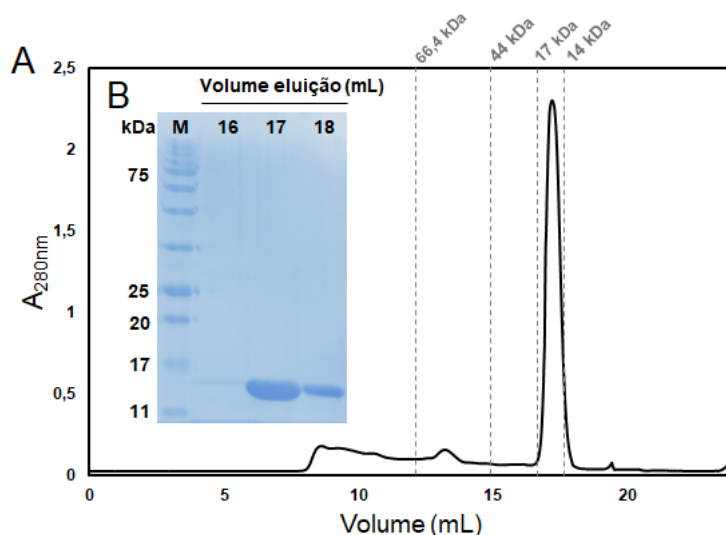
**Figura 4.7- SDS-PAGE (12%) referentes à clivagem enzimática do domínio C-SH2 com a protease HRV-3C e com as frações provenientes do segundo passo cromatográfico, através da IMAC. (A)-** SDS PAGE (12%) com as frações obtidas antes (poço 1) e depois (poço 2) da digestão com a protease HRV-3C. **(B)-** M- Marcador de pesos moleculares NZYColour Protein Marker II; 1- Fração Flow-through; 2 a 5 – Frações de 5 mL de lavagem; 6 a 8- Frações de 5 mL de eluição.

Nesta segunda cromatografia, a proteína de interesse foi eluída na fase de lavagem com o tampão de ligação (poços 1 a 5), uma vez que a proteína deixou de apresentar afinidade para os íons de níquel, devido à remoção His<sub>6</sub>-tag por ação da protease HRV-3C. Pela análise do gel, também é possível verificar que as *tags* Trx-His<sub>6</sub>, bem como outras proteínas contaminantes, permaneceram adsorvidas à matriz da coluna, tendo sido eluídas aquando da adição de imidazol ao tampão, por este competir com as histidinas pela ligação aos íons de níquel (poços 6 a 8).

Contudo, nos poços 1 a 5 é possível verificar presença de bandas correspondentes a outras proteínas, para além da banda correspondente ao peso molecular do domínio C-SH2 de SHP-1.

Tendo em conta que a utilização de proteína com elevado grau de pureza é determinante para a realização dos ensaios pretendidos, procedeu-se a uma cromatografia de exclusão

molecular, como passo final de purificação. Este método permite a separação eficaz de proteínas com diferentes pesos moleculares. Foi utilizada uma coluna Superdex 200™ 10/300 GL e o cromatograma obtido encontra-se apresentado na figura 4.8A, que indica a existência de um pico principal, eluído aproximadamente entre os 16 e os 18 mL de eluente. As frações correspondentes ao pico obtido, foram analisadas por SDS-PAGE (figura 4.8B). Todos os outros picos de menor intensidade do cromatograma correspondem a proteínas contaminantes, já antes observados no gel da figura 4.7B.



**Figura 4.8- Cromatograma e SDS-PAGE (12%) obtidos no último passo de purificação do domínio C SH2, realizado por cromatografia de exclusão molecular. (A)-**  $A_{280nm}$  em função do volume de eluição. O pico observado corresponde à eluição da proteína de interesse (entre os 16 e 18 mL). As linhas a cinzento representam o máximo dos picos das proteínas eluídas aquando da realização da calibração da coluna. **(B)-** SDS-PAGE (12%) com as frações correspondentes ao pico obtido na eluição. M- Marcador de pesos moleculares NZYColour Protein Marker II; 16 a 18- Frações eluídas.

O gel confirma que o pico corresponde à eluição da proteína de interesse, pois verifica-se a presença da banda correspondente ao peso molecular da mesma (14,4 kDa). Utilizando a calibração da coluna realizada com proteínas de peso molecular conhecido, foi possível verificar o estado oligomérico do domínio de interesse, demonstrando o mesmo ser monomérico.

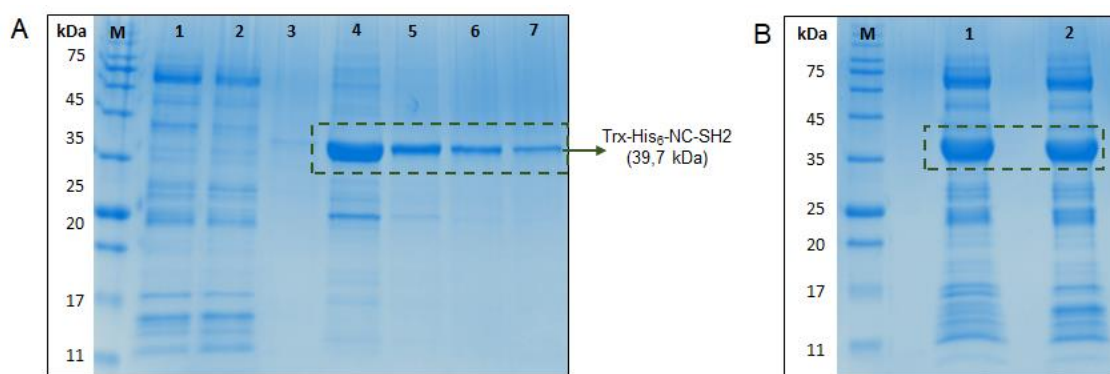
#### 4.1.2.4 Expressão e tentativas de purificação dos domínios N-SH2 e NC-SH2 de SHP-1

A expressão e as tentativas de purificação das diferentes construções dos domínios N-SH2 e NC-SH2 de SHP-1 seguiram os mesmos passos descritos anteriormente para o domínio C-SH2. Isto é, depois das células recolhidas terem sofrido lise celular e subsequente ultracentrifugação, as proteínas de interesse contidas na fração solúvel foram purificadas por cromatografia de afinidade.



Numa primeira abordagem as construções mais simples destes domínios no vetor de expressão pHTP8 (Trx-His<sub>6</sub>-N-SH2 e Trx-His<sub>6</sub>-NC-SH2, consultar tabela 3 da secção 3.1.2.1.) foram purificadas utilizando novamente a coluna Strep-TactinXT™ (5 mL), seguindo o mesmo princípio descrito anteriormente. Este passo de purificação foi analisado através de SDS-PAGE, apresentado na figura 4.9A. De notar que nesta secção apenas estão apresentados os géis obtidos na tentativa de purificação do domínio NC-SH2, pois a purificação do domínio N-SH2 foi realizada do mesmo modo e obtiveram-se resultados similares (géis da purificação do domínio N-SH2 presentes em anexo na figura 7.2). Através da análise do gel da figura 4.9A é possível confirmar, nos poços 4, 5, 6 e 7, referentes à eluição, a presença do domínio de interesse fundido com as *tags* Trx e His<sub>6</sub> na região N-terminal (39,7 kDa).

O passo seguinte à cromatografia realizada, que removeu a maioria dos contaminantes, seria cortar as *tags* Trx e His<sub>6</sub> das proteínas de interesse através de uma digestão com a protease HRV-3C. Esta clivagem enzimática aconteceu de igual forma e nas mesmas condições descritas anteriormente para o domínio C-SH2. A digestão foi analisada através de SDS-PAGE (figura 4.9B), o qual demonstrou a ineficiência do processo, uma vez que a banda correspondente ao peso molecular do domínio NC-SH2 fundido com as *tags* (39,7 kDa) continua a ser observada ao final da digestão.



**Figura 4.9- SDS-PAGE (12%) da tentativa de purificação da construção Trx-His<sub>6</sub>-NC-SH2. (A)- SDS-PAGE (12%) com as frações provenientes da cromatografia de afinidade realizada utilizando a coluna Strep-TactinXT™. M- Marcador de pesos moleculares NZYColour Protein Marker II; 1- Fração Flow-through; 2- Fração de lavagem; 3 a 7- Frações de 5 mL de eluição; (B)- SDS-PAGE (12%) com as frações obtidas antes (poço 1) e depois (poço 2) da digestão com a protease HRV-3C.**

Atendendo ao descrito, repetiu-se o mesmo passo cromatográfico para as mesmas construções, desta vez tentando efetuar a clivagem enzimática diretamente na coluna, tal como já havia sido realizado com C-Tir. As frações recolhidas foram analisadas por SDS-PAGE, verificando-se uma vez mais a ineficiência do corte.

Da mesma maneira, tentou-se purificar os domínios por IMAC, utilizando a coluna HisTrap™ (5 mL), onde tal como descrito anteriormente, a eluição dos mesmos seria conseguida através da adição de imidazol ao tampão. Nas diferentes construções, a His<sub>6</sub>-tag encontrava-se na proximidade da região reconhecida pela protease HRV-3C. Desta forma, a adsorção das proteínas de fusão à coluna, poderia facilitar o corte das *tags* Trx-His<sub>6</sub> dos domínios de interesse diretamente na mesma, pois essa região estaria teoricamente mais acessível para a clivagem

enzimática. Uma vez mais as frações recolhidas foram analisadas por SDS-PAGE, que indicou que o corte permaneceu sem êxito (figura 7.3, em anexo).

O insucesso da clivagem enzimática deveu-se provavelmente ao facto do local de reconhecimento pela protease HRV-3C estar inacessível. Tal poderia estar a acontecer por três razões: **i)** a proteína Tiorredoxina fundida aos domínios de interesse podia estar a interagir com os mesmos, bloqueando estereoquimicamente esse local; **ii)** a conformação tridimensional dos domínios de interesse poderia bloquear intrinsecamente o local de reconhecimento; **iii)** os domínios de interesse expressos na fração solúvel serem, na verdade, agregados solúveis.

De forma a verificar e controlar o cenário descrito no ponto **i)** foram idealizadas construções dos domínios de interesse no vetor pHTP10, que contém NusA em vez de Trx. Assim sendo, essas construções (NusA-His<sub>6</sub>-N-SH2 e NusA-His<sub>6</sub>-NC-SH2, consultar tabela 3 da secção 3.1.2.1.) foram expressas e purificadas, da mesma maneira descrita anteriormente. Repetiram-se os passos de purificação descritos e verificou-se uma vez mais o insucesso da clivagem enzimática (figura 7.4, em anexos).

Visto que o problema não estava na *tag* de solubilidade escolhida, foi decidido desenhar novas construções, adicionando um *linker* de 8 aminoácidos entre o local de reconhecimento da protease HRV-3C e os domínios de interesse. Desta forma, o possível bloqueio do local de reconhecimento da protease causado pela conformação dos domínios de interesse seria atenuado ou deixaria de existir, pois a região de reconhecimento passaria a estar mais distante da estrutura tridimensional desses domínios (ponto **ii)**). A expressão e purificação dessas construções (Trx-His<sub>6</sub>-L8-N-SH2, Trx-His<sub>6</sub>-L8-NC-SH2) (consultar tabela 3 da secção 3.1.2.1.) foram realizadas tal como descrito anteriormente. De novo, a purificação dos domínios de interesse foi realizada utilizando a coluna Strep-TactinXT™ (5 mL) e efetuando o corte fora da coluna. Contudo, o processo de clivagem manteve-se ineficaz (figura 7.5, em anexos).

Outra abordagem passou pela alteração da protease responsável pela clivagem enzimática. Foram idealizadas novas construções no vetor pHTP8, trocando a região de reconhecimento da protease HRV-3C pela região reconhecida pela protease TEV. Uma vez mais, repetindo a expressão e purificação dessas construções (Trx-His<sub>6</sub>-TEV-L8-N-SH2 e Trx-His<sub>6</sub>-TEV-NC-SH2, consultar tabela 3 da secção 3.1.2.1.), os mesmos problemas persistiram ao tentar realizar a clivagem enzimática.

A última solução passaria por expressar e purificar estes domínios no vector de expressão pCoofy18, que não contém *tags* de solubilidade, com o intuito da purificação dos domínios de interesse não passar pela digestão com proteases (His<sub>10</sub>-N-SH2 e His<sub>10</sub>-NC-SH2, consultar tabela 3 da secção 3.1.2.1.). Contudo, no tempo em que foi realizada a presente dissertação, nunca se conseguiu obter a expressão desses domínios no vector pCoofy18 na fração solúvel, pelo que essas construções foram abandonadas.

## 4.2. Caracterização de C-Tir através da espectroscopia de RMN

A fosforilação das tirosinas presentes nas regiões ITIM de C-Tir é um requisito para a sua interação com domínios SH2 (descrito em detalhe na secção 1.1.2.1.). Dessa forma, o estudo da



fosforilação e dos seus efeitos estruturais em C-Tir é fundamental para o entendimento do processo de recrutamento e interação com as proteínas do hospedeiro.

A espectroscopia de RMN desempenha um papel crucial na caracterização de proteínas desordenadas (como C-Tir) e nas interações em que as mesmas estão envolvidas, pois é a única técnica espectroscópica que fornece informações estruturais e dinâmicas a nível atômico deste tipo de proteínas. Da mesma maneira, esta técnica permite a monitorização de modificações pós-traducionais em proteínas, como a fosforilação, uma vez que os sinais de ressonância das mesmas são muito sensíveis às mudanças induzidas por essas modificações. Desta feita, a espectroscopia de RMN demonstra ser a técnica mais adequada ao estudo pretendido.

A fosforilação de C-Tir e a atribuição das ressonâncias da proteína nesse estado fosforilado foi o primeiro passo desta fase do trabalho. Uma vez alcançados esses objetivos, estávamos em condições para estudar a interação entre esta e o domínio C-SH2 de SHP-1.

#### 4.2.1 Atribuição das ressonâncias de C-Tir fosforilada

A atribuição das ressonâncias da cadeia principal de aminoácidos da proteína C-Tir não-fosforilada já tinha sido realizada anteriormente no laboratório (*Vieira et al, artigo em preparação*).

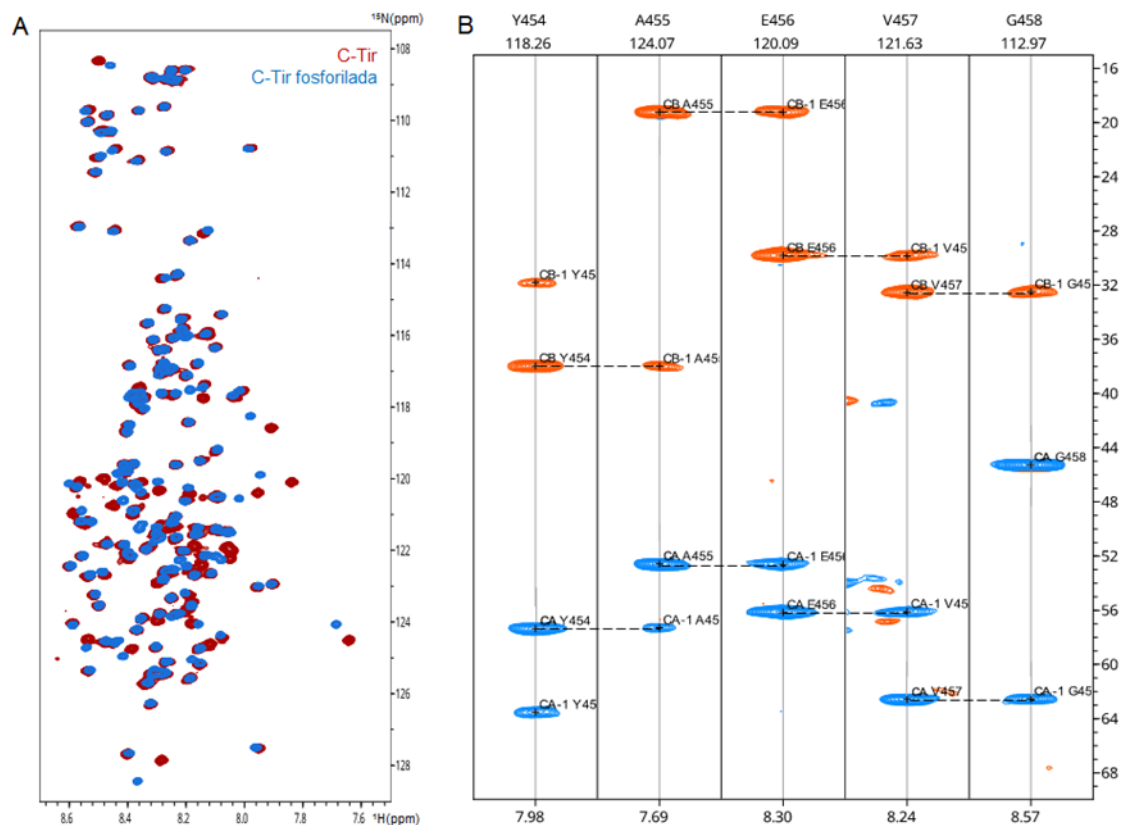
Os ensaios de fosforilação indicaram que apesar da fosforilação ser específica para os resíduos de tirosinas, esta modificação provoca alterações no espectro de C-Tir (figura 4.10A). Embora a maioria dos desvios químicos não sejam afetados, existem alguns que o são, nomeadamente as tirosinas fosforiladas e os resíduos vizinhos tornando a atribuição das novas ressonâncias um desafio. Para os resíduos de serina (pSer) e treonina fosforilados (pThr), são normalmente observadas fortes alterações nos desvios químicos das amidas da cadeia principal em comparação com as versões não-fosforiladas. Em contraste, a fosforilação da tirosina (pTyr) não induz um grande deslocamento na posição do resíduo modificado, provavelmente devido à posição mais distal do grupo hidroxilo modificado na cadeia lateral em relação à cadeia principal (consultar a secção 1.2.2.1. para mais detalhe). Foi notado que em C-Tir, certos resíduos contíguos a pTyr são mais afetados que a própria tirosina, o que indica que os desvios químicos dos resíduos perto de uma pTyr podem ser usados para identificar a fosforilação de resíduos de tirosina por espectroscopia de RMN.

Para identificar as ressonâncias da proteína fosforilada de forma inequívoca, procedemos à reatribuição das ressonâncias de C-Tir fosforilada, recorrendo a experiências multidimensionais (2D,3D) de correlação entre  $^1\text{H}$ ,  $^{13}\text{C}$  e  $^{15}\text{N}$ .

Dessa forma, a amostra de C-Tir duplamente marcada com  $^{15}\text{N}$  e  $^{13}\text{C}$  foi fosforilada pela cinase Fyn durante a noite, a 25 °C e foram adquiridas as experiências 2D HSQC e as experiências 3D HNCO e HNCACB.

A experiência 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC correlaciona o azoto (N) do grupo amida com o protão (H) que lhe está diretamente ligado. Com base nesta experiência, foram realizadas experiências 3D para atribuir as ressonâncias de RMN à sequência da proteína fosforilada.

Atendendo a que a atribuição das ressonâncias da proteína não-fosforilada já tinha sido feita foi possível realizar a reatribuição sequencial das ressonâncias da proteína fosforilada, utilizando uma das experiências 3D mais úteis e comuns para esse propósito, a experiência 3D HNCACB. O espectro desta experiência contém correlações inter e intraresiduais, uma vez que correlaciona o NH e o N de um resíduo (*i*) com os C $\alpha$  e C $\beta$  do próprio resíduo e do resíduo anterior (*i*-1).



**Figura 4.10- Atribuição das ressonâncias de C-Tir fosforilada. (A)-** Sobreposição dos espectros 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC de C-Tir (vermelho) e C-Tir fosforilada (azul) em 20 mM HEPES pH 6.8, 150 mM NaCl, 12 mM MgCl<sub>2</sub> e 5.5 mM ATP. **(B)-** Faixa da atribuição sequencial dos resíduos Y454, A455, E456, V457 e G458 da sequência de C-Tir fosforilada, alcançada através da experiência 3D HNCACB. Os C $\alpha$  (fase negativa) estão representados a azul e a cor-de-laranja estão representados os C $\beta$  (fase positiva). Os tracejados a cinzento indicam a atribuição sequencial dos átomos referidos (entre os C $\alpha$  e o C $\alpha$ -1/ C $\beta$  e C $\beta$ -1).

Extrapolando do espectro HSQC temos as experiências 3D representadas num cubo, onde o plano xy contém a frequência das correlações NH e a terceira dimensão corresponde à frequência dos carbonos. A extração de planos perpendiculares ao eixo  $^{15}\text{N}$  permitiu a conexão dos resíduos (*i*) com os seus resíduos adjacentes (*i*-1). Este procedimento foi auxiliado pelo conhecimento dos padrões de dispersão específicos dos desvios químicos dos carbonos. A faixa escolhida e apresentada na figura 4.10B, demonstra a atribuição sequencial dos aminoácidos Y454, A455, E456, V457 e G458 da sequência de C-Tir fosforilada. As ressonâncias de fase negativa correspondem aos C $\alpha$  e as de fase oposta aos C $\beta$ . Esta experiência envolve a transferência de magnetização entre os  $^{13}\text{C}\alpha$  e  $^{15}\text{N}$  através das constantes de acoplamento J ( $J_{\text{NC}\alpha}$ ) da cadeia principal da proteína. De forma geral,  $2J_{\text{NC}\alpha}$  são ligeiramente menores do que

$1J_{NC\alpha}$ , tornando os sinais intra-resíduo normalmente mais fortes do que os sinais de correlações entre resíduos.<sup>74</sup> Assim sendo e como é visível na figura 4.10, conseguimos distinguir os  $C\alpha$  dos  $C\alpha-1$  e os  $C\beta$  dos  $C\beta-1$ , uma vez que os sinais do próprio resíduo são mais intensos do que os do resíduo anterior. Contudo, se tal não for perceptível, podem cometer-se erros na atribuição das ressonâncias.

A experiência adquirida é uma versão equivalente da experiência 3D CBCANH (descrita anteriormente na secção 1.2.2.) que explora as mesmas conectividades, mas que é mais sensível, uma vez que é independente do  $J_{HC}$ . Usualmente esta experiência é adquirida juntamente com a experiência 3D HN(CO)CACB para estabelecer a correspondência de forma inequívoca dos  $C\alpha$  e  $C\beta$  e dos  $C\alpha-1$  e  $C\beta-1$ , uma vez que essa experiência apenas correlaciona o  $C\alpha$  e  $C\beta$  do resíduo anterior ( $i-1$ ) com o NH e o N do próprio resíduo ( $i$ ). Além disso, a experiência 3D HNCA, que correlaciona o HN e o N de um resíduo ( $i$ ) com o  $C\alpha$  do próprio resíduo e resíduo anterior ( $i-1$ ), poderia ter sido utilizada, em conjunto com HN(CO)CA, que correlaciona apenas o HN e o N de um resíduo ( $i$ ) com o  $C\alpha$  do resíduo anterior ( $i-1$ ), para atribuir os  $C\alpha$  de alguns resíduos específicos da proteína, uma vez que estas experiências são bastante sensíveis e permitem aceder a resoluções espectrais superiores. Contudo, o facto da atribuição das ressonâncias de C-Tir não-fosforilada já ter sido realizada, tornou possível a reatribuição das mesmas, apenas com a aquisição da experiência 3D descrita (HNCACB).

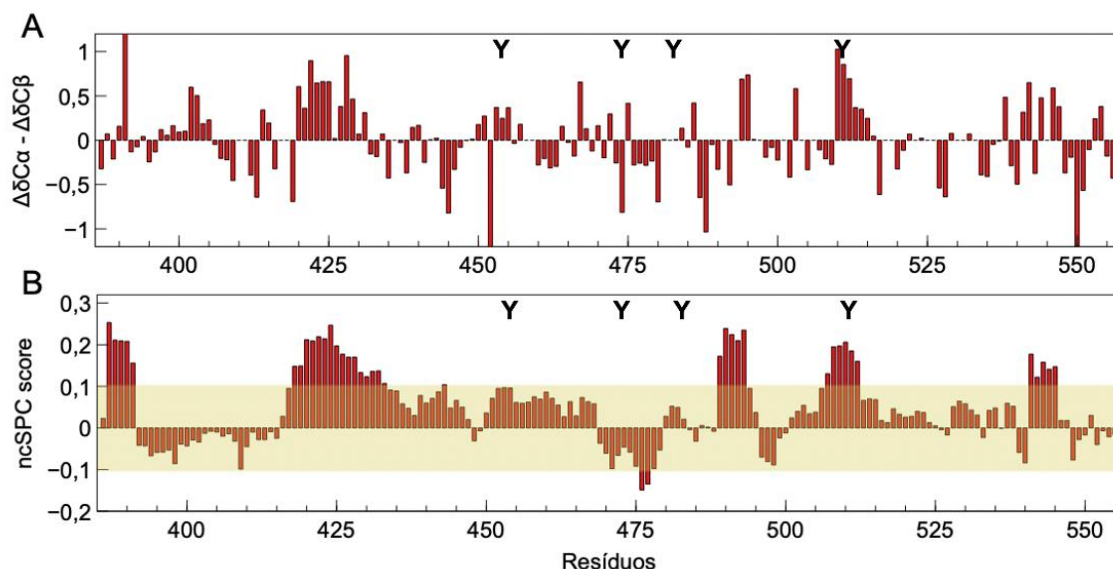
A experiência 3D mais sensível de RMN, 3D HNCO, foi utilizada para atribuir o carbono do grupo carbonilo (CO), pois esta correlaciona o HN e o N de um resíduo ( $i$ ) com o CO do resíduo anterior ( $i-1$ ).

Conseguiram-se atribuir os desvios químicos  $C\alpha$  e CO de 173 e 160 resíduos, respetivamente, 161 desvios químicos de azoto e hidrogénio dos grupos amida e 150 desvios químicos  $C\beta$  (considerando que as glicinas não têm  $C\beta$ ), de um total de 173 resíduos. Não se conseguiu atribuir os desvios químicos de azoto e hidrogénio das 12 prolinas, pois estes resíduos não têm o protão da amida ligado ao azoto na ligação peptídica e, por essa razão não aparecem no espectro HSQC. De notar ainda que foram atribuídos 160 desvios químicos CO, pois para esse propósito só foi adquirida a experiência 3D HNCO, fazendo com que os resíduos anteriores às prolinas, não conseguissem ter os seus desvios químicos atribuídos, pela razão referida anteriormente. Para tal, poderíamos ter recorrido à experiência 3D HN(CA)CO que correlaciona o HN e o N de um resíduo ( $i$ ) com o próprio CO e com o CO do resíduo anterior ( $i-1$ ).

#### 4.2.2 Propensão de elementos de estrutura secundária de C-Tir fosforilada

Os desvios químicos são bons indicadores da conformação estrutural de uma proteína. As suas alterações em relação a valores dos mesmos em conformação *random coil* dão informação acerca da propensão de elementos de estrutura secundária. Essas alterações são designadas por desvios químicos secundários ( $\Delta\delta$ ).<sup>76</sup>

Conforme descrito em maior detalhe na seção 1.2.2., valores positivos de desvios químicos secundários do C $\alpha$  ( $\Delta\delta^{C\alpha}$ ) indicam propensão para hélices- $\alpha$ , enquanto que valores negativos indicam a preferência para folhas- $\beta$ . Para os valores de desvios químicos secundários do C $\beta$  ( $\Delta\delta^{C\beta}$ ) o efeito é o inverso. Portanto, é conveniente monitorizar a diferença  $\Delta\delta^{C\alpha} - \Delta\delta^{C\beta}$ , estando esses resultados apresentados na figura 4.11A e os  $\Delta\delta^{C\alpha}$  e  $\Delta\delta^{C\beta}$  encontram-se apresentados em anexos (figura 7.6). De notar que na presente dissertação calculámos os desvios químicos secundários utilizando desvios químicos em conformação *random coil* corrigidos e ajustados para IDPs.<sup>82</sup>



**Figura 4.11- Desvios químicos secundários ( $\Delta\delta$ ) calculados para C-Tir no estado fosforilado.** Os  $\Delta\delta^{C\alpha}$  e  $\Delta\delta^{C\beta}$  foram calculados através da diferença entre os desvios químicos C $\alpha$  e C $\beta$  experimentais e os desvios químicos referentes aos mesmos átomos corrigidos para IDPs. A posição das tirosinas fosforiladas na sequência de C-Tir encontram-se representadas pela letra Y. **(A)-** Diferença entre  $\Delta\delta^{C\alpha}$  e  $\Delta\delta^{C\beta}$ ; **(B)-** Score calculado através do método ncSPC que resulta da combinação dos desvios químicos secundários N, HN, CO, C $\alpha$  e C $\beta$ . Valores entre -0,1 e 0,1 foram considerados como propensões *random-coil* (área representada a amarelo).

Analisando os dados da figura 4.11A, observamos duas regiões com valores positivos de diferença  $\Delta\delta^{C\alpha} - \Delta\delta^{C\beta}$  que indicam a propensão para estruturas hélice- $\alpha$ . A primeira região está situada entre os resíduos D420 e R432 e a segunda está localizada numa das regiões ITIM mimetizadas por C-Tir, compreendida entre os resíduos T510 e A515.

Os desvios químicos secundários foram calculados considerando os resíduos individualmente, mas também usando o método ncSPC, que considera a influência dos resíduos vizinhos de cada resíduo, melhorando a eficiência e confiabilidade das previsões de propensão dos elementos de estrutura secundária. Este método combina não só os  $\Delta\delta^{C\alpha}$  e  $\Delta\delta^{C\beta}$ , mas também os  $\Delta\delta^{CO}$ ,  $\Delta\delta^N$  e  $\Delta\delta^{HN}$ , permitindo identificar e corroborar essas mesmas previsões. Os resultados obtidos através do método descrito estão apresentados na figura 4.11B.

Através da sua análise, conseguem-se observar várias regiões com propensão para estruturas de hélice- $\alpha$ . As duas regiões observadas anteriormente no cálculo dos  $\Delta\delta^{C\alpha} - \Delta\delta^{C\beta}$ ,

mantêm a mesma propensão e apresentam-se com maior definição. Essas regiões são conhecidas por serem responsáveis por interagir com proteínas do hospedeiro, especialmente a segunda região, que envolve um dos ITIMs mimetizados por C-Tir e que compreende o resíduo Y511 fosforilado. Para além das regiões mencionadas, verifica-se a presença de uma nova região com a mesma propensão para estrutura helicoidal, situada entre os resíduos Q487 e G491. É de notar que estas regiões são contíguas a sítios de fosforilação por cinases do hospedeiro, como o resíduo de serina S434, e a já mencionada tirosina Y483.<sup>114</sup> Possivelmente, a estruturação que é sugerida para essa zona depois da mesma ser fosforilada, é uma característica importante e que afeta a capacidade de interação e controlo do hospedeiro.

O efector Tir quando é libertado no citosol da célula hospedeira, migra para a membrana plasmática através de um processo que permanece desconhecido até aos dias de hoje (consultar secção 1.1.2. para mais detalhe). Regiões ITIMs e ITAMs, já vem sido descritas como capaz de interagir com a membrana e adotar conformações helicoidais.<sup>115,116</sup> Uma interpretação possível destes resultados pode sugerir que a propensão de hélice- $\alpha$  destas regiões ITIMs, para além de ser importante para o reconhecimento de proteínas do hospedeiro, pode ser importante para o reconhecimento da membrana, tendo a fosforilação um papel importante na regulação desta interação. Estudos adicionais, explorando a interação de C-Tir fosforilada e não-fosforilada com membranas, serão necessários para sustentar essa hipótese. Por último, observamos uma região adicional, entre os resíduos R388 e P392, com tendência para formar hélice- $\alpha$ . Assumindo uma topologia em *hairpin* (consultar a figura 1.1 presente na secção 1.1.2.), esta região está a jusante da região transmembranar, muito perto da membrana plasmática. Assim sendo, a hélice- $\alpha$  que é sugerida para essa região de C-Tir poderá ter um papel importante na estabilização de Tir na membrana, ou fazer possivelmente ainda parte do domínio intramembranar de Tir, que está previsto adotar uma estrutura secundária hélice- $\alpha$ .

#### **4.2.3 Reação de fosforilação de C-Tir *in vitro* monitorizada em tempo real descontínuo**

Conforme já foi referido anteriormente, assume-se que a fosforilação é um processo obrigatório para a interação entre C-Tir e os domínios SH2. Este facto associado à influência que esta modificação pós-traducional parece ter na capacidade de recrutamento de proteínas do hospedeiro, faz com que o estudo da fosforilação seja essencial para a compreensão de todo este processo.

De facto, a dinâmica e os papéis funcionais dos eventos de fosforilação individuais dos resíduos de tirosinas presentes nas regiões ITIM consenso de C-Tir não são compreendidos. Tendo isto em consideração, foi realizada uma reação de fosforilação *in vitro* monitorizada por espectroscopia de RMN em tempo real descontínuo, com o intuito de estudar a cinética e a extensão da fosforilação das tirosinas. Tal como descrito anteriormente (consultar secção 1.2.2.2.) este método baseia-se na recolha de alíquotas a diferentes tempos de reação após a

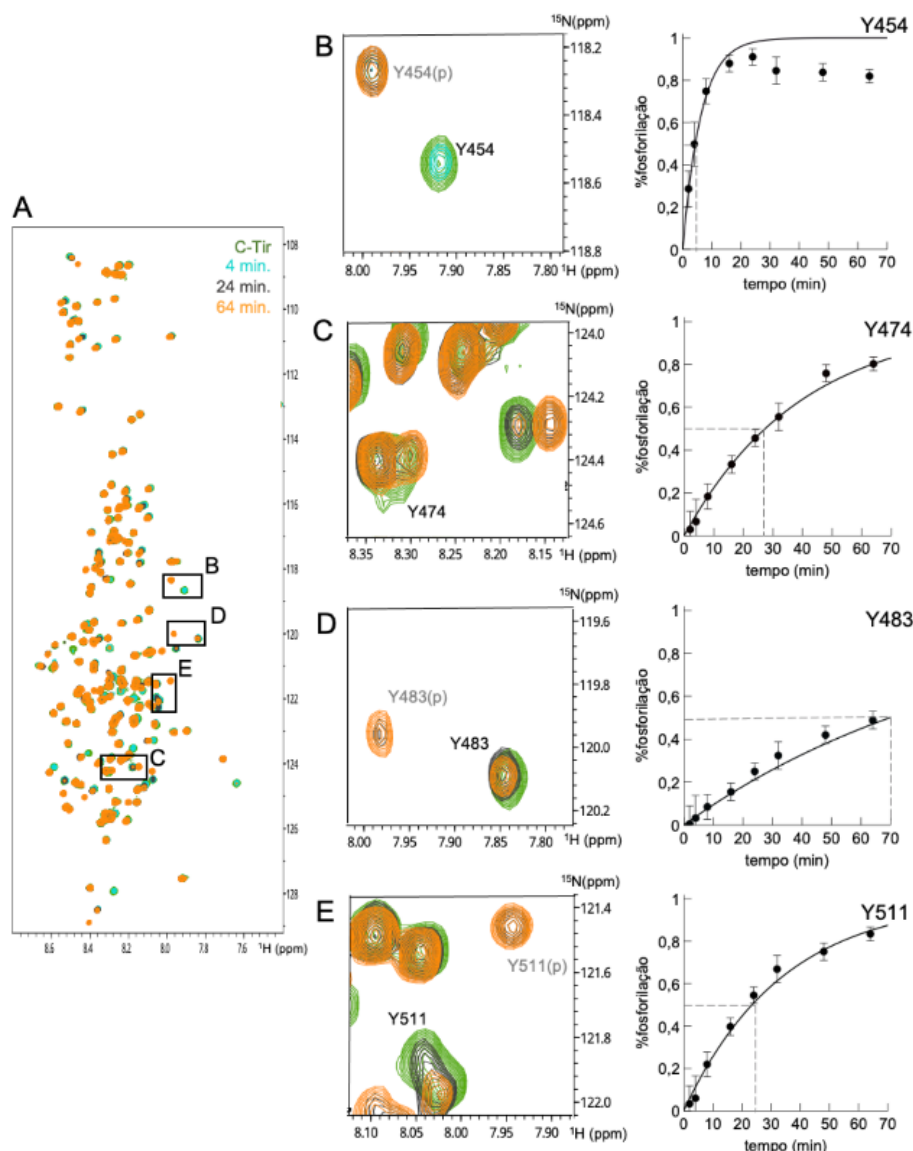
adição da cinase à proteína de interesse. A monitorização e quantificação dos níveis de fosforilação individual das quatro tirosinas foi alcançada através da aquisição de espectros 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC aos diferentes tempos de reação escolhidos.

Assim sendo, a amostra de C-Tir foi incubada com a cinase Fyn. Contrariamente ao que acontece em todos os outros ensaios realizados na presença de ATP, nesta situação a reação foi realizada na presença de apenas 2 mM de ATP de forma a tornar possível a análise da seletividade da cinase Fyn com o objetivo de identificar os potenciais comportamentos cinéticos das quatro tirosinas de C-Tir.

Nos espectros consecutivos 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC aos diferentes tempos de reação, aparecem três tipos de sinais: i) os que não mudam durante o decurso da reação; ii) os que desaparecem; e iii) os que aparecem (figura 4.12). Os sinais do primeiro tipo pertencem a aminoácidos que, tendo em conta a sequência e a falta de estrutura terciária da proteína, se encontram longe dos locais de fosforilação e que por essa razão não são afetados por esse evento. Estes são vistos como sinais de controlo de RMN cujas intensidades determinam um fator de escala ( $F_{\text{escala}}$ ) usado para corrigir pequenas diferenças nas condições experimentais (consultar secção **3.2.2.2.**). Os sinais que desaparecem e/ou aparecem pertencem aos resíduos de tirosinas que sofreram fosforilação ou a resíduos que estão próximos destes locais de fosforilação e que sofrem alterações nos seus ambientes químicos.

Graças à atribuição das ressonâncias da proteína no estado fosforilado, o nível de fosforilação de cada tirosina pôde ser monitorizado e quantificado pela intensidade relativa dos seus sinais ao longo da reação. As intensidades relativas e a percentagem de fosforilação de cada resíduo foram calculadas de acordo com as equações 3.1 e 3.2, apresentadas na secção **3.2.2.2.** Estas informações quantitativas foram ajustadas a curvas de acumulação específicas desses locais, curvas essas que refletem os comportamentos de fosforilação das tirosinas individualmente. O ajuste das curvas foi alcançado através da equação 3.3 (secção **3.2.2.2.**), estando as mesmas apresentadas na figura 4.12B-E.

A ressonância da tirosina Y474 fosforilada (figura 4.12C) moveu-se para uma região do espectro com muitos picos sobrepostos, tornando difícil a sua identificação. Contudo, o descrito não foi um problema uma vez que os níveis de fosforilação individuais foram calculados indiretamente a partir das ressonâncias referentes à população de cada tirosina que permanece não-fosforilada, segundo as equações já mencionadas.



**Figura 4.12- Reação de fosforilação de C-Tir in vitro monitorizado por RMN em tempo real descontínuo.** (A)- Espetro 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC de C-Tir (verde) e após 4 (azul ciano), 24 (cinzento) e 64 minutos de reação (cor de laranja) em 20 mM HEPES pH 6.8, 150 mM NaCl, 12 mM  $\text{MgCl}_2$  e 5.5 mM ATP. (B-E)- Esquerda: regiões dos espectros 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC que contêm as ressonâncias dos resíduos Y454, Y474, Y483 e Y511 aos diferentes tempos de reação referidos. Direita: curvas de fosforilação que contêm informação dos níveis de fosforilação individuais dos resíduos Y454, Y474, Y483 e Y511 após 2, 4, 8, 16, 24, 32, 48 e 64 minutos de reação. As barras de erro indicam o desvio padrão do  $F_{\text{escala}}$  usado para corrigir possíveis diferenças nas condições experimentais.

As experiências de RMN permitiram revelar que C-Tir pode ser fosforilada simultaneamente em diferentes posições e que cada fosforilação ocorreu de forma eficiente, mas com velocidades aparentemente diferentes. Pela análise do perfil das curvas da fosforilação é possível verificar que a tirosina Y454 é a primeira a ser fosforilada, alcançado praticamente 50 % de fosforilação passado 4 minutos de reação. Por outro lado, as restantes tirosinas: Y511, Y474 e Y483 alcançam o mesmo nível de fosforilação depois de 24, 32 e 64 minutos, respetivamente.

A cinética de fosforilação do resíduo Y483 parece ser a mais lenta dos resíduos de tirosina, como pode ser observado através do declive menos acentuado da sua curva de fosforilação, em comparação com as restantes. Neste resíduo, no último tempo de reação medido, atinge-se apenas aproximadamente 50% de fosforilação, enquanto que os outros resíduos de tirosina estão aproximadamente 90% fosforilados, ao mesmo tempo de reação. Dessa maneira, a velocidade das fosforilações individuais pode ser classificada pela ordem: Y454>Y511>Y474>Y483.

Através da figura 4.12, podemos ainda verificar que a localização das tirosinas na sequência de C-Tir tem influência na seletividade e afinidade da cinase Fyn, pois apesar das quatro tirosinas serem fosforiladas, o padrão de fosforilação de cada uma delas é distinto.

Para simplificar a análise, definimos os quatro locais de fosforilação em três classes: rápida (Y454), intermédia (Y474 e Y511) e lenta (Y483). Num sistema sequencial de fosforilação, a ordem pela qual os locais de fosforilação são ativados pela ação da cinase determina os estados de fosforilação.<sup>117</sup> Contudo, não há evidências de que a fosforilação estudada se trata de um sistema sequencial, em que a fosforilação de um resíduo dependeria dos outros (i.e., fosforilação condicional), uma vez que a fosforilação nas quatro tirosinas ocorre em simultâneo, com velocidades diferentes (i.e., fosforilação preferencial). Para corroborar o descrito, seria necessário proceder a mutações nestes resíduos fosforilados. Por exemplo, uma mutação na tirosina mais rápida (Y454) deveria aumentar a velocidade de fosforilação das tirosinas definidas como intermédias e lenta.

Na sinalização celular, os eventos de fosforilação e desfosforilação atuam como um interruptor molecular, pelo que a dinâmica/velocidade de fosforilação pode ter implicações funcionais importantes. A fosforilação da tirosina Y454, desempenha um papel menor e menos eficiente na polimerização de actina, interagindo diretamente com a proteína N-WASP, sem a necessidade do recrutamento da proteína adaptadora Nck (consultar secção 1.1.2.1. para maior detalhe). Este resíduo, para além de fazer parte de um ITIM mimetizado por Tir, faz igualmente parte de um motivo NPY, conhecido por interagir com proteínas com domínio I-BAR, que por sua vez, são reconhecidos por induzir a polimerização de actina.<sup>118</sup> Em Tir-EHEC, sabe-se que a formação dos pedestais de actina depende da interação desta tirosina não-fosforilada com o domínio I-BAR da proteína IRSp53 (consultar secção 1.1.2.1.). Tendo em conta que as sequências de Tir-EPEC e EHEC são conservadas, é provável que a mesma interação se verifique em EPEC, ainda que a mesma não tenha sido documentada. Desta forma, embora não seja evidente a justificação de uma cinética mais rápida nesta posição, podemos especular acerca do possível papel deste resíduo na regulação da atividade de Tir, funcionando a fosforilação como um interruptor molecular que influencia a seletividade do resíduo para as diferentes interações referidas. Contrariamente à tirosina Y454, a tirosina Y474 tem um papel bem estabelecido na formação dos pedestais de actina. Isto associado às considerações descritas anteriormente permite sugerir que as tirosinas Y511 e Y483 sejam os resíduos responsáveis pelo recrutamento dos domínios SH2 de SHP-1. Dessa forma, a cinética mais



favorável na tirosina Y511 poderá indicar uma interação preferencial por este resíduo em detrimento da tirosina Y483.

Ao abrigar quatro locais de fosforilação de resíduos de tirosina, C-Tir expande consideravelmente o repertório da sua regulação combinatória. Em comparação com a fosforilação de um único resíduo, a fosforilação em múltiplos locais (*multisite phosphorylation*) aumenta consideravelmente as possibilidades de regulação da função da proteína. Na verdade, uma proteína com N locais de fosforilação pode existir em  $2^N$  estados de fosforilação e cada um desses estados pode ter uma característica funcional diferente.<sup>117,119</sup> Assim sendo, C-Tir pode ter um máximo de 16 estados de fosforilação de tirosinas ( $2^4$ ). Como a reação de fosforilação estudada é realizada *in vitro*, não estamos a considerar o envolvimento de fosfatases e de outros fatores que existem dentro da célula e por isso, é a ação da cinase Fyn e a sua especificidade para cada tirosina que define quais os estados e as populações relativas de fosforilação que temos num dado instante. Por exemplo, baseando-nos nestes dados, aos 2 minutos de reação nestas condições, a probabilidade de existir apenas o estado mono-fosforilado de C-Tir é elevada e o resíduo fosforilado seria a Y454, uma vez que esta é a que apresenta uma cinética mais rápida e a única a ter população fosforilada nesse instante. O mesmo raciocínio pode ser feito para os restantes tempos reacionais.

Conforme descrito no capítulo inicial (secção 1.2.2.2.), para além deste método, monitorização em tempo real descontínuo, existe outro que poderia ser usado para estudar a reação de fosforilação por RMN: monitorização em tempo real contínuo. A escolha do método usado revelou-se favorável pois de outra forma não conseguiríamos detetar o primeiro ponto de medição definido (2 minutos) pois, na melhor situação, os espetros de RMN 2D demorariam 4 minutos a ser adquiridos. Da mesma maneira, com o método escolhido, garantimos que os tempos de medição definidos correspondem de facto aos tempos reacionais. Isto é particularmente importante nos resíduos onde a cinética é rápida para que os níveis de fosforilação obtidos estejam sujeitos ao mínimo de erro possível. Exemplo disto é o resíduo Y454, cuja cinética de fosforilação é bastante rápida e 50% da fosforilação é alcançada ao fim de 4 minutos. Se a monitorização da fosforilação deste resíduo fosse feita de forma contínua, a análise estaria provavelmente comprometida, pois os tempos de medição definidos não corresponderiam à realidade.

#### 4.2.4 Interação entre C-Tir e o domínio C-SH2 de SHP-1

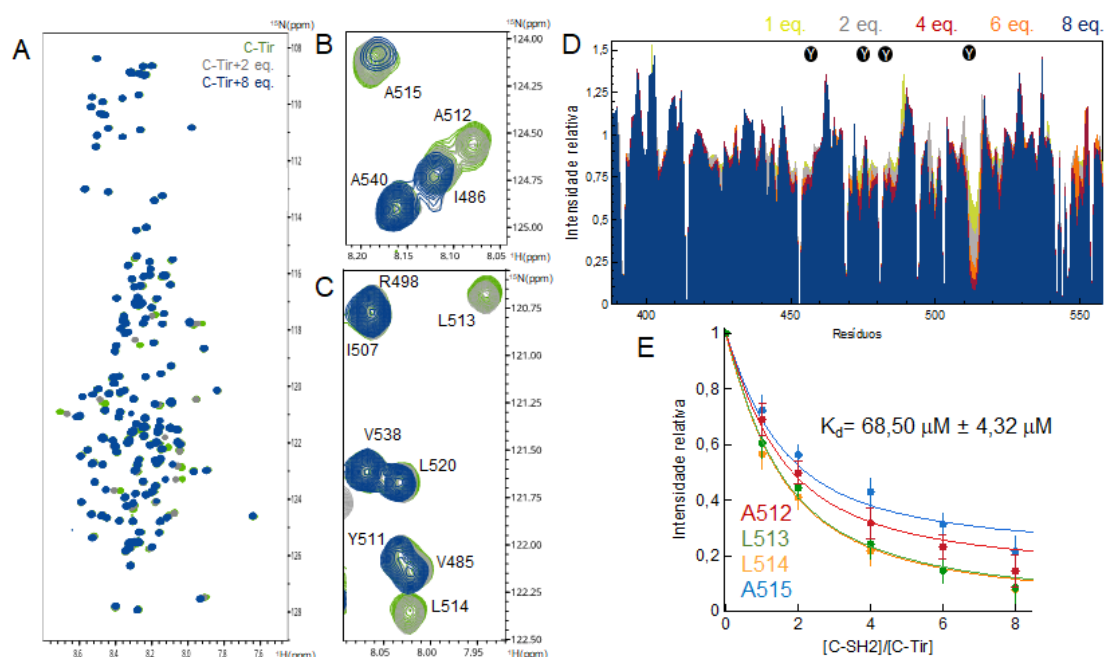
Atendendo ao descrito na literatura, C-Tir necessita de ser fosforilada nos seus resíduos de tirosina para interagir com domínios SH2 (consultar a secção 1.1.2.1.). Para estudar se intrinsecamente C-Tir tem a capacidade de interagir o domínio C-SH2 de SHP-1, ou se essa interação só consegue ser estabelecida quando C-Tir se encontra no estado fosforilado, foram realizados ensaios de interação entre as proteínas não-fosforiladas.

Recorremos à espectroscopia de RMN para testar esta interação. C-Tir isotopicamente marcada com  $^{15}\text{N}$  foi titulada com diferentes equivalentes do domínio C-SH2 e foram medidos

espectros 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC. A atribuição das ressonâncias da cadeia principal de C-Tir realizada anteriormente no laboratório foi utilizada para monitorizar perturbações nas intensidades observadas nos espectros 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC após adição do domínio C-SH2. Os resultados destes ensaios de interação encontram-se apresentados na figura 4.13.

A primeira observação que se consegue fazer através dos espectros sobrepostos da figura 4.13A é que C-Tir mantém o seu carácter desordenado mesmo na presença do domínio C-SH2, uma vez que os desvios químicos dos prótons das amidas dos aminoácidos da cadeia principal continuam a exibir baixa dispersão, agrupando-se na região entre os 7.6 e os 8.6 ppm.

Os espectros de C-Tir titulada com o domínio C-SH2 de SHP-1 aos equivalentes referidos, demonstram uma redução sistemática de intensidades das ressonâncias dos resíduos situados nas proximidades da região ITIM que compreende a tirosina Y511, como apresentado na figura 4.13D. Esses resíduos são a A512, L513, L514 e A515, que participam na interface da interação com o domínio C-SH2 (figura 4.13B e 4.13C).



**Figura 4.13- Estudo da interação entre C-Tir e o domínio C-SH2 de SHP-1. (A)-** Sobreposição dos espectros  $^1\text{H}$ - $^{15}\text{N}$  HSQC de C-Tir isotopicamente marcada com  $^{15}\text{N}$  com quantidades crescentes do domínio C-SH2, em 20 mM HEPES pH 6.8 e 150 mM NaCl. A verde está representada C-Tir, a cinzento C-Tir com 2 eq. de C-SH2 e a azul com 8 eq. do mesmo domínio. **(B-C)-** Regiões dos espectros  $^1\text{H}$ - $^{15}\text{N}$  HSQC sobrepostos que contém os resíduos A512, L513, L514 e A515 envolvidos na interação. **(D)-** Gráfico das intensidades relativas dos resíduos de C-Tir ao longo da titulação (1 eq. representado a amarelo, 2 eq. a cinzento, 4 eq. a vermelho, 6 eq. a cor de laranja e 8 eq. a azul). **(E)-** Gráfico das intensidades relativas dos quatro resíduos localizados na interface da interação em função de  $[\text{C-SH2}]/[\text{C-Tir}]$ . O valor de  $K_d$  apresentado representa a força de interação entre os resíduos mencionados de C-Tir e o domínio C-SH2 de SHP-1. As barras de erro indicam o desvio padrão do  $F_{\text{escala}}$  usado para corrigir possíveis diferenças nas condições experimentais.

Pela análise das figuras 4.13A-C, é difícil prever qual o regime de permuta entre o estado livre e complexado de C-Tir com o domínio C-SH2, uma vez que apenas se verifica a diminuição sistemática da intensidade de algumas ressonâncias. No entanto, com base no  $K_d$  obtido ( $\mu\text{M}$ ) e como não são visíveis alterações nos desvios químicos entre o estado livre e complexado da proteína, podemos prever que nos encontramos entre o regime de permuta lenta e intermédia, ainda que as ressonâncias dos resíduos relativos ao estado complexado da proteína, não sejam observáveis.

Por si, o domínio C-SH2 é relativamente pequeno (14,4 kDa) contudo, à baixa temperatura a que as experiências foram adquiridas ( $10^\circ\text{C}$ ) o aumento da viscosidade é promovido e a proteína comporta-se como uma entidade maior, devido ao aumento do tempo de correlação rotacional ( $\tau_c$ ) ( $\tau_c$  ( $25^\circ\text{C}$ ) = 5,55 ns;  $\tau_c$  ( $10^\circ\text{C}$ ) = 8,54 ns, calculado através das equações 7.1 e 7.2 presentes em anexo).<sup>120,121</sup> C-Tir ao interagir com o domínio referido, altera as suas propriedades hidrodinâmicas e comporta-se como uma molécula maior. Quanto maior for o tamanho da molécula, maior será o efeito do alargamento dos sinais, ou dito de outra forma, menor será a sua intensidade. Assim sendo, a redução de intensidade relativa dos sinais observada é indicativa da diminuição da mobilidade induzida pela interação destes resíduos de C-Tir com o domínio C-SH2 (maior  $\tau_c$ ). Outro fator que promove o alargamento dos sinais está associado a fenómenos de permuta química que juntamente com o aumento de tamanho favorecem o desaparecimento dos sinais dos resíduos diretamente envolvidos na interação. Em estudos de interação entre IDPs e parceiros globulares, foi observado o alargamento dos sinais abaixo do nível de deteção, mesmo em condições saturantes, sugerindo uma interação transitória e dinâmica, caracterizada por mais de dois estados.<sup>122-124</sup>

Devido à natureza transitória da interação e às propriedades hidrodinâmicas das partes ordenadas e desordenadas do complexo formado serem distintas, a espectroscopia de RMN geralmente deteta apenas as regiões que permanecem flexíveis e que não estão envolvidas diretamente na interação.

Na figura 4.13E encontra-se apresentado o gráfico das intensidades relativas dos resíduos alterados em função do rácio molar  $[\text{C-SH2}]/[\text{C-Tir}]$ . Foi calculado o valor de  $K_d$  para obter um valor mensurável e discutir a força de interação dessa região com o domínio referido. Através de um ajuste global (conseguido com a equação 3.4, apresentada na secção 3.2.3.2.) obteve-se um valor de  $K_d$  igual a  $68,5 \mu\text{M} \pm 4,32 \mu\text{M}$ .

Face aos resultados obtidos, a fosforilação parece ser de facto um requisito para a interação entre as tirosinas presentes nas regiões ITIM de C-Tir e o domínio C-SH2 de SHP-1. Contudo, intrinsecamente C-Tir aparenta ter algumas características estruturais que em certas condições podem favorecer esta interação nas proximidades da região ITIM que contém a Y511, sem a necessidade da mesma estar fosforilada.

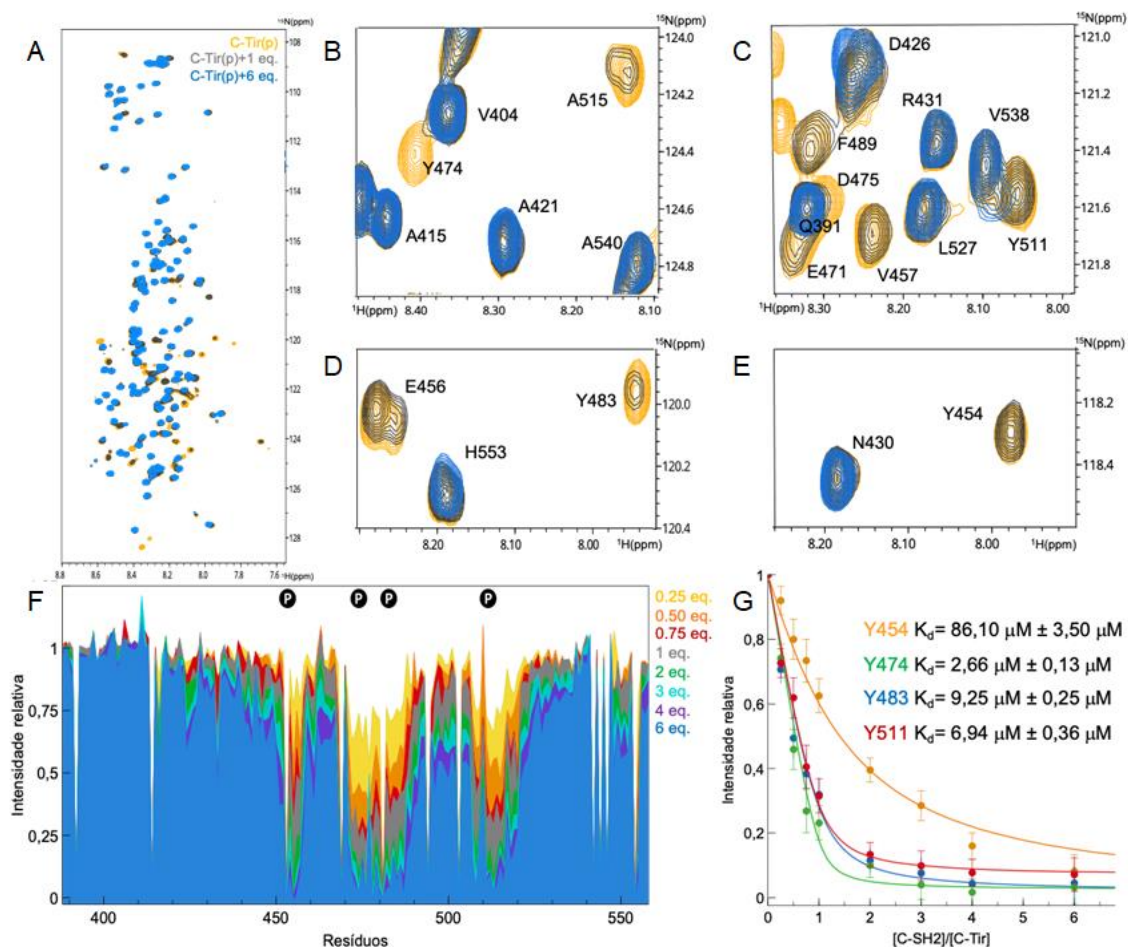
#### 4.2.5 Interação entre C-Tir fosforilada e o domínio C-SH2 de SHP-1

Recorreu-se novamente à espectroscopia de RMN para testar a interação entre C-Tir no estado fosforilado e o domínio C-SH2 de SHP-1. Para tal, C-Tir isotopicamente marcada com  $^{15}\text{N}$  foi fosforilada pela ação da cinase Fyn e foi titulada com diferentes equivalentes do domínio C-SH2 de SHP-1. Foram medidos espectros 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC para cada ponto da titulação. A reatribuição das ressonâncias da cadeia principal de C-Tir fosforilada foi utilizada para monitorizar perturbações nas intensidades dos espectros 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC após adição do domínio C-SH2. Os resultados destes ensaios de interação encontram-se apresentados na figura 4.14.

Conforme conseguimos observar através dos espectros adquiridos em cada ponto da titulação (figura 4.14A-E), a reação de fosforilação foi completa, pois não se verifica a presença de picos duplos nos resíduos fosforilados nem nos seus resíduos vizinhos. De igual forma, é possível verificar que C-Tir na presença do domínio C-SH2 mantém as suas características desordenadas, uma vez que os desvios químicos dos prótons das amidas dos resíduos da cadeia principal continuam a exibir baixa dispersão.

Podemos verificar mudanças subtis nos espectros de C-Tir fosforilada na presença de concentrações crescentes de C-SH2 (figura 4.14A). Contudo, tal como já havia acontecido na interação entre C-Tir não-fosforilada e o domínio C-SH2 (secção 4.2.4.), é visível a diminuição sistemática das intensidades das ressonâncias de alguns resíduos, indo algumas para além dos limites de deteção da técnica. Uma vez mais, a diminuição da intensidade é indicativa dos resíduos de C-Tir que participam na interface desta interação. As considerações referidas acerca da viscosidade, do tempo de correlação rotacional ( $\tau_c$ ) e de permuta química no ensaio anterior, são também aplicadas neste ensaio para justificar os espectros obtidos e a diminuição de intensidades dos resíduos diretamente envolvidos na interação.

Pela análise da figura 4.14F, na qual estão representadas as intensidades relativas dos resíduos ao longo da titulação, conseguimos verificar que, ao contrário do ensaio entre as proteínas não-fosforiladas, onde a interação parece ocorrer de forma bastante específica e localizada, neste ensaio a interação parece afetar mais regiões da sequência de C-Tir. A mesma figura, demonstra que essas regiões são os quatro ITIMs mimetizados por C-Tir, estando os resíduos compreendidos nas mesmas diretamente envolvidos na interação com o domínio C-SH2.



**Figura 4.14- Estudo da interação entre C-Tir fosforilada e o domínio C-SH2 de SHP-1. (A)-** Sobreposição dos espectros  $^1\text{H}$ - $^{15}\text{N}$  HSQC de C-Tir isotopicamente marcada com  $^{15}\text{N}$  com quantidades crescentes do domínio C-SH2, em 20 mM HEPES pH 6.8, 150 mM NaCl, 12 mM  $\text{MgCl}_2$  e 5.5 mM ATP. A amarelo está representada C-Tir fosforilada, a cinzento C-Tir com 1 eq. de C-SH2 e a azul com 6 eq. do mesmo domínio. **(B-E)-** Regiões dos mesmos espectros  $^1\text{H}$ - $^{15}\text{N}$  HSQC sobrepostos que contém os resíduos Y474, Y511, Y483 e Y454 envolvidos na interação. **(F)-** Gráfico das intensidades relativas dos resíduos de C-Tir ao longo da titulação. **(G)-** Gráfico das intensidades relativas dos quatro resíduos de tirosinas em função de  $[\text{C-SH2}]/[\text{C-Tir}]$ . Os valores de  $K_d$  apresentados são referentes a cada resíduo de tirosina, sendo que foram calculados tendo em conta os seus resíduos vizinhos que fazem parte das regiões ITIM de C-Tir. As barras de erro indicam o desvio padrão do  $F_{\text{escala}}$  usado para corrigir possíveis diferenças nas condições experimentais.

A medição da força de interação de cada tirosina fosforilada, cada uma compreendida na sua região ITIM, para o domínio C-SH2, foi obtida através do cálculo do valor de  $K_d$ . Na figura 4.14G encontra-se apresentado o gráfico das intensidades relativas de cada tirosina em função do rácio molar  $[\text{C-SH2}]/[\text{C-Tir}]$ . Procedemos a um ajuste global (usando a equação 3.4 da secção 3.2.3.2.) para cada um destes resíduos, no qual foram tidos em conta também os resíduos vizinhos de cada tirosina, que estão compreendidos nas regiões ITIM e que são igualmente afetados pela interação. Os gráficos das intensidades relativas destes resíduos vizinhos que estão

compreendidos nas regiões ITIM em função do rácio molar [C-SH2/C-Tir] encontram-se apresentadas na figura 7.7, em anexos.

Tendo em conta os papéis que as tirosinas Y454 e Y474 têm na formação dos pedestais de actina, era de esperar que a interação acontecesse preferencialmente através dos ITIMs que compreendem as tirosinas Y483 e Y511. Contudo, analisando os resultados obtidos, é possível verificar que apesar da afinidade do domínio C-SH2 ser menor para a região que compreende a tirosina Y454 ( $K_d = 86,1 \mu\text{M} \pm 3,50 \mu\text{M}$ ), o mesmo não se verifica para a região da tirosina Y474, que apresenta uma afinidade semelhante para o mesmo domínio quando comparada com as regiões ITIM que compreendem as tirosinas Y483 e Y511 ( $K_d = 2,66 \mu\text{M} \pm 0,13 \mu\text{M}$ ;  $9,25 \mu\text{M} \pm 0,25 \mu\text{M}$ ;  $6,94 \mu\text{M} \pm 0,36 \mu\text{M}$ , respetivamente). De notar que neste ensaio estamos apenas a trabalhar com uma versão minimalista do sistema, uma vez que apenas estamos a analisar a interação do domínio C-SH2 com o efector Tir. É expectável que ao realizar o ensaio no contexto do domínio *tandem* (NC-SH2), a seletividade das regiões ITIM seja melhorada, pois estaremos na presença dos dois domínios SH2 de SHP-1, capazes de interagir com Tir, com especificidades diferentes.

Comparando os valores obtidos para a interação entre C-Tir no estado não-fosforilado e no estado fosforilado com o domínio C-SH2, conseguimos observar uma afinidade dez vezes maior para a região ITIM que compreende o resíduo Y511, quando esse resíduo está fosforilado. De relembrar que esta era a única região que conseguia interagir intrinsecamente com o domínio C-SH2, sem que C-Tir estivesse fosforilada. Esta informação permite corroborar o que já havia sido descrito na secção anterior acerca da importância da fosforilação na interação abordada.

De uma forma mais correta, os valores de  $K_d$  obtidos são valores de  $K_d$  aparentes, pois são os valores que conseguimos retirar assumindo um modelo de interação 1:1, mas onde existe competição interna pela interação. C-Tir contém quatro tirosinas que são fosforiladas e que competem internamente pela ligação ao domínio C-SH2. Se não estivéssemos nestas condições, seria de esperar um valor de  $K_d$  mais favorável para cada local de interação individual.

## **5. Conclusões e perspectivas futuras**





O equilíbrio dos mecanismos patogénicos bacterianos *versus* a resposta imune do hospedeiro determina se a bactéria será eliminada pelo hospedeiro ou se a bactéria sobreviverá e se multiplicará no mesmo. As bactérias desenvolveram uma miríade de maneiras pelas quais podem inclinar a balança a seu favor. Vários patógenos bacterianos usam sistemas de secreção do tipo III (T3SS) para modular o comportamento da célula hospedeira. O T3SS canaliza diretamente as proteínas efetoras das bactérias para as células eucarióticas, onde provoca uma regulação disfuncional das proteínas de sinalização presentes nas mesmas e subverte os seus processos normais, aumentando assim a capacidade da bactéria de colonizar o hospedeiro. O presente trabalho visou contribuir para um conhecimento mais amplo dos mecanismos de virulência e manipulação do hospedeiro por bactérias ao estudar a fosforilação da região C-terminal da proteína efetora Tir (C-Tir) e a sua interação com um domínio SH2 de uma proteína do hospedeiro. C-Tir mimetiza quatro regiões endógenas de reconhecimento molecular do hospedeiro, conhecidos por ITIMs, que compreendem cada um uma tirosina, fosforilada por cinases do hospedeiro. Dessa forma, Tir adquire a capacidade de recrutar várias proteínas com domínios SH2, com destaque para a fosfatase SHP-1, uma proteína importante em algumas vias de sinalização da resposta imune. A interação entre as duas, suprime a produção de citocinas pro-inflamatórias, que iriam, em condições normais, inibir a multiplicação bacteriana. Vários estudos indicam que a fosforilação de Tir é um requisito para esta interação. Assim sendo, propusemos reconstruir a versão fosforilada de Tir *in vitro* a fim de estudar os detalhes moleculares da interação entre C-Tir e SHP-1 e perceber como molecularmente a fosforilação dos resíduos de tirosina molda esta interação.

Em primeiro lugar, a concretização desse objetivo dependia do sucesso da sobre-expressão e purificação de C-Tir e dos domínios SH2 de SHP-1. Contrariamente ao que aconteceu com os domínios N-SH2 e NC-SH2, onde se verificaram problemas na purificação, a expressão e purificação de C-Tir e do domínio C-SH2 foi alcançada, tendo o trabalho avançado com estas duas proteínas.

Foi desenvolvido um protocolo com sucesso para obter a versão fosforilada de C-Tir, por ação de uma cinase de tirosinas do hospedeiro (Fyn) e com recurso a experiências multidimensionais de RMN, conseguimos reatribuir as ressonâncias dos desvios químicos da proteína no estado fosforilado, a partir de uma atribuição que já havia sido realizada anteriormente no laboratório. Uma vez alcançada a atribuição das ressonâncias, estávamos em condições de avançar com os estudos subsequentes propostos. Da análise dos desvios químicos secundários foram identificadas regiões com propensão para estruturas helicoidais, nomeadamente regiões localizadas perto dos resíduos S434, Y483 e Y511, passíveis de sofrerem fosforilação por cinases do hospedeiro e conhecidas por interagirem com proteínas do mesmo, quando os resíduos mencionados são fosforilados. Especulamos que a dita estruturação poderá ter implicações no recrutamento de proteínas do hospedeiro ou na interação com membranas biológicas. A fosforilação poderá eventualmente minimizar a associação com membranas, tornando o efector Tir mais acessível para sequestrar proteínas alvo e modular o comportamento da célula hospedeira.

Dada a importância da fosforilação na interação entre o efector Tir e as proteínas do hospedeiro, procurámos estudar a cinética e a extensão da reação de fosforilação das quatro tirosinas compreendidas nas regiões ITIM mimetizadas. Para tal, estabelecemos um método de monitorização em tempo real descontínuo da reação de fosforilação *in vitro*, que designámos dessa forma por serem feitas medições de RMN de diferentes alíquotas, aos diferentes tempos de extensão reacional. A quantificação dos níveis de fosforilação dos locais individuais permitiu revelar que Tir é capaz de sofrer fosforilação em múltiplos locais simultaneamente, podendo coexistir as quatro tirosinas fosforiladas. Contudo, as cinéticas e extensões das quatro posições são diferentes: enquanto que a tirosina Y483 é a que sofre uma fosforilação mais lenta, a tirosina Y454 é a que é mais rapidamente fosforilada, estando as tirosinas Y511 e Y474 numa posição intermédia no que toca à velocidade de fosforilação.

Atendendo ao descrito na literatura, a fosforilação das tirosinas das regiões ITIMs apresentava-se como um requisito para a interação entre C-Tir e os domínios SH2 das proteínas do hospedeiro. Nos ensaios de interação realizados entre a proteína C-Tir não-fosforilada com o domínio C-SH2, verificou-se que a interação pode ocorrer, estando a região ITIM que compreende a tirosina Y511 na interface da mesma. Como descrito anteriormente, essa região apresenta propensão para adotar estrutura hélice- $\alpha$ , o que nos leva a concluir que intrinsecamente C-Tir aparenta ter algumas características estruturais, que em determinadas condições podem favorecer esta interação.

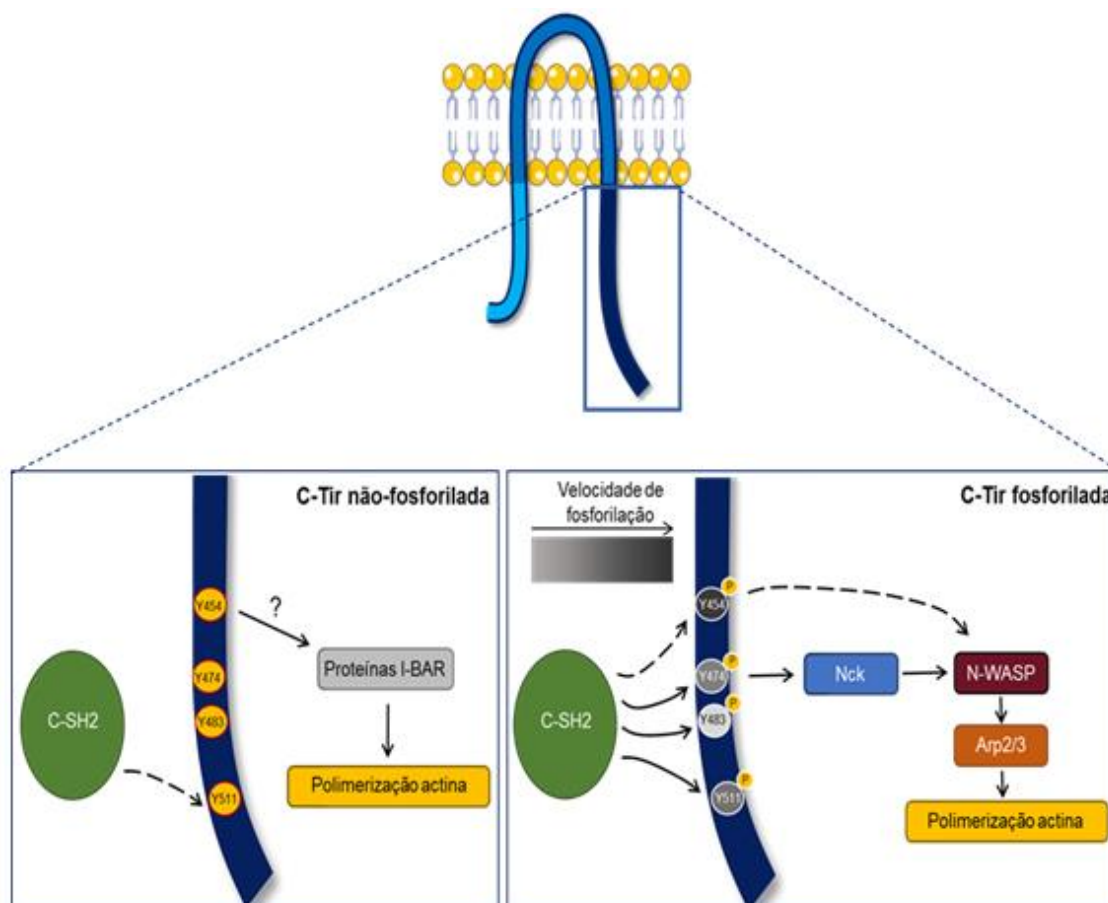
Atendendo ao papel que as tirosinas Y454 e Y474 têm na formação dos pedestais de actina, o cenário mais provável seria que as regiões que compreendem as restantes tirosinas, a Y483 e a Y511 fossem as responsáveis pelo recrutamento dos domínios SH2 de SHP-1. Dessa maneira, e a partir da observação do facto da cinética da tirosina Y511 ser mais favorável que a da tirosina Y483, teríamos uma indicação sobre uma possível interação preferencial do domínio C-SH2 com o primeiro resíduo em detrimento do segundo. Contudo, os ensaios entre C-Tir fosforilada e o domínio C-SH2 não demonstraram o sugerido, uma vez que apesar dos mesmos corroborarem que as regiões ITIM são as responsáveis pela interação, esta é bastante promíscua, uma vez que o domínio C-SH2 se demonstra capaz de interagir com as tirosinas com afinidades muito semelhantes. A exceção é a interação com a região ITIM que compreende a tirosina Y454, onde a força de interação é menor comparativamente com as restantes. A promiscuidade de ligação verificada, explica-se, em parte, por estarmos a trabalhar com uma versão simplista do sistema proposto, pois na presente dissertação apenas foi realizada a análise da interação do domínio C-SH2 com o efector Tir. Assim sendo, o cenário verificado pode ser alterado se o ensaio for repetido no contexto do domínio *tandem* (NC-SH2), onde é expectável que a seletividade das regiões ITIM seja melhorada, pois estaremos na presença dos dois domínios SH2 da fosfatase SHP-1, que teoricamente serão ambos capazes de interagir com Tir, com especificidades diferentes.

A afinidade da interação da região ITIM que contém a tirosina Y511 fosforilada é dez vezes maior que a afinidade da mesma região quando o mesmo resíduo não está fosforilado. Desta maneira, apesar da observação referida acerca da possível interação entre as proteínas

não-fosforiladas descrever uma novidade no âmbito do conhecimento científico, consegue-se perceber a importância da fosforilação na interação estudada.

As diferentes cinéticas e extensões da reação de fosforilação podem ter implicações na sinalização celular, uma vez que os eventos de fosforilação/desfosforilação atuam na célula como um interruptor molecular, muitas vezes essencial para definir consequências funcionais. Conforme descrito, a fosforilação da tirosina Y454 desempenha um papel menor e pouco eficiente na formação dos pedestais de actina, ao estar envolvido num mecanismo pouco eficiente, que é independente da proteína Nck para a formação dos mesmos. Para além do descrito, este resíduo está presente no motivo NPY conservado em patógenos que provocam lesões A/E. Em Tir-EHEC, quando o resíduo não está fosforilado é reconhecido por proteínas com domínios I-BAR, importantes para a polimerização de actina. Graças à conservação que existe entre as sequências de Tir-EPEC e EHEC, podemos sugerir que as mesmas interações possam existir em EPEC, embora estas não estejam documentadas e comprovadas. Nos ensaios realizados foi verificado que a tirosina Y454 é a que é mais rapidamente fosforilada, sendo capaz de interagir com o domínio C-SH2, embora apresente menos afinidade para o mesmo, comparativamente com as restantes tirosinas. Assim sendo, a possível explicação para a fosforilação tão rápida deste resíduo é que o mesmo pode funcionar como um interruptor molecular capaz de influenciar e regular a seletividade relativa deste resíduo para a interação com domínios SH2 (quando fosforilado) ou para os domínios I-BAR (quando desfosforilado). Desta maneira, esta tirosina aparenta ter um papel fundamental na regulação da atividade de Tir na célula hospedeira. Estudos adicionais, explorando a mutação deste resíduo seriam necessários para sustentar o descrito, pois com a mutação proposta, conseguiríamos verificar a importância deste resíduo nos mecanismos descritos e conseguiríamos verificar se existe uma correlação sequencial de fosforilação entre as restantes tirosinas fosforiladas.

As principais conclusões referidas anteriormente, retiradas a partir do trabalho desenvolvido na presente dissertação encontram-se representadas na figura 5.1. Em suma, este trabalho abre caminho para novos estudos que visam a compreensão de mecanismos de virulência com implicações na saúde e na doença. Contudo, na presente dissertação, devido aos problemas existentes na purificação dos domínios N-SH2 e NC-SH2 acabamos por trabalhar com uma versão minimalista do sistema proposto. No futuro, propomos aumentar a complexidade do mesmo, analisando a interação de C-Tir também com os domínios N-SH2 e NC-SH2. Ao fazê-lo, conseguiríamos estudar a direcionabilidade e a especificidade de cada domínio e poderíamos construir uma imagem mais complexa e perceptível da interação entre C-Tir e SHP-1, conhecendo os determinantes moleculares que estão na base desta interação e na ativação desta fosfatase por esta proteína efetora, essencial à infeção bacteriana.



**Figura 5.1- Representação esquemática das conclusões retiradas a partir do trabalho desenvolvido na presente dissertação.** À esquerda está representada C-Tir-EPEC em situação de não-fosforilação, em que a interação com o domínio C-SH2 é capaz de ocorrer através da região ITIM mimetizada por Tir que contém a tirosina Y511. À direita encontra-se representada C-Tir-EPEC quando a mesma é fosforilada por cinases do hospedeiro. As quatro tirosinas são fosforiladas simultaneamente, a diferentes velocidades, estando essas velocidades representadas numa escala de cinza (aumento do mais claro para o mais escuro). Nesta situação, C-Tir interage com o domínio C-SH2 através das quatro tirosinas compreendidas nas quatro regiões consenso ITIM mimetizadas, com afinidades semelhantes entre as tirosinas, menos para a tirosina Y454, cuja afinidade para o domínio é menor. Este resíduo é provavelmente importante na regulação da atividade de Tir na célula hospedeira, uma vez que é o resíduo que sofre uma fosforilação mais rápida e que ao mesmo tempo é capaz de interagir com o domínio C-SH2 (com menos afinidade) ou com a proteína N-WASP (num mecanismo indireto e pouco eficiente que conduz à formação dos pedestais de actina) quando o efector bacteriano se encontra fosforilado. Em Tir-EHEC, quando o mesmo se encontra desfosforilado, ao fazer também parte do motivo NPY, é capaz de interagir com proteínas com domínios I-BAR, que induzem a polimerização de actina. A mesma interação é sugerida para Tir-EPEC.

## **6. Bibliografia**



1. World Health Organization. Public Health Review of the Enterohaemorrhagic Escherichia Coli Outbreak in Germany. *Reg Off Eur World Heal Organ*. 2011;(June).
2. Afset JE, Bergh K, Bevanger L. High prevalence of atypical enteropathogenic Escherichia coli (EPEC) in Norwegian children with diarrhoea. *J Med Microbiol*. 2003;52(11):1015-1019. doi:10.1099/jmm.0.05287-0
3. Donnenberg MS, Kaper JB, Finlay BB. Interactions between enteropathogenic Escherichia coli and host epithelial cells. *Trends Microbiol*. 1997;5(3):109-114. doi:10.1016/S0966-842X(97)01000-7
4. Kenny B, DeVinney R, Stein M, Reinscheid DJ, Frey EA, Finlay BB. Enteropathogenic E. coli (EPEC) transfers its receptor for intimate adherence into mammalian cells. *Cell*. 1997;91(4):511-520. doi:10.1016/S0092-8674(00)80437-7
5. Pearson JS, Giogha C, Lung T, Hartland LE. The genetics of Enteropathogenic Escherichia coli Virulence. *Annual Review of Genetics*. 2016;50:493-513
6. Donnenberg MS, Whittam TS. Pathogenesis and evolution of virulence in enteropathogenic and enterohemorrhagic Escherichia coli. *J Clin Invest*. 2001;107(5):539-548. doi:10.1172/JCI12404
7. Yan D, Wang X, Luo L, Cao X, Ge B. Inhibition of TLR signaling by a bacterial protein containing immunoreceptor tyrosine-based inhibitory motifs. *Nat Immunol*. 2012;13(11):1063-1071. doi:10.1038/ni.2417
8. Clements A, Young JC, Constantinou N, Frankel G. Infection strategies of enteric pathogenic Escherichia coli. *Gut Microbes*. 2012;3(2):71-87. doi:10.4161/gmic.19182
9. Raman R, Thomas RG, Weiner MW, et al. Diversity of the Human Intestinal Microbial Flora. *Science (80- )*. 2005;308(5728):1635-1638.
10. Kotloff KL, Nataro JP, Blackwelder WC, et al. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): A prospective, case-control study. *Lancet*. 2013;382(9888):209-222. doi:10.1016/S0140-6736(13)60844-2
11. Gaytán MO, Martínez-Santos VI, Soto E, González-Pedrajo B. Type three secretion system in attaching and effacing pathogens. *Front Cell Infect Microbiol*. 2016;6(OCT):1-25. doi:10.3389/fcimb.2016.00129
12. Milon A, Oswald E, De Rycke J. Rabbit EPEC: A model for the study of enteropathogenic Escherichia coli. *Vet Res*. 1999;30(2-3):203-219. doi:10.1016/S0928-4249(99)80017-0
13. Sudlow M. O157 H7 Hus. *Nejm*. 1971;64(12):1237-1238.
14. Jarvis KG, Girón JA, Jerse AE, Mcdaniel TK, Donnenberg MS, Kaper JB. Enteropathogenic Escherichia coli contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. *Proc Natl Acad Sci U S A*. 1995;92(17):7996-8000. doi:10.1073/pnas.92.17.7996
15. Mota LJ, Cornelis GR. The bacterial injection kit: Type III secretion systems. *Ann Med*. 2005;37(4):234-249. doi:10.1080/07853890510037329
16. Xu L, Liu Y. Protein secretion systems in bacterial pathogens. *Front Biol (Beijing)*.

- 2014;9(6):437-447. doi:10.1007/s11515-014-1333-z
17. Kubori T, Matsushima Y, Nakamura D, et al. Supramolecular structure of the salmonella typhimurium type III protein secretion system. *Science* (80- ). 1998;280(5363):602-605. doi:10.1126/science.280.5363.602
  18. Vallance BA, Finlay BB. Exploitation of host cells by enteropathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A*. 2000;97(16):8799-8806. doi:10.1073/pnas.97.16.8799
  19. Mcdaniel TK, Jarvis KG, Donnenberg MS, Kaper JB. A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc Natl Acad Sci U S A*. 1995;92(5):1664-1668. doi:10.1073/pnas.92.5.1664
  20. Elliott SJ, Sperandio V, Giron JA, et al. The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infect Immun*. 2000;68(11):6115-6126. doi:10.1128/IAI.68.11.6115-6126.2000
  21. Daniell SJ, Takahashi N, Wilson R, et al. The filamentous type III secretion translocon of enteropathogenic *Escherichia coli*. *Cell Microbiol*. 2001;3(12):865-871. doi:10.1046/j.1462-5822.2001.00168.x
  22. Iguchi A, Thomson NR, Ogura Y, et al. Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. *J Bacteriol*. 2009;91(1):347-354. doi:10.1128/JB.01238-08
  23. Tobe T, Beatson SA, Taniguchi H, et al. An extensive repertoire of type III secretion effectors in *Escherichia coli* O157 and the role of lambdoid phages in their dissemination. *Proc Natl Acad Sci U S A*. 2006;103(40):14941-14946. doi:10.1073/pnas.0604891103
  24. Via A, Uyar B, Brun C, Zanzoni A. How pathogens use linear motifs to perturb host cell networks. *Trends Biochem Sci*. 2015;40(1):36-48. doi:10.1016/j.tibs.2014.11.001
  25. Selbach M, Paul FE, Brandt S, et al. Host Cell Interactome of Tyrosine-Phosphorylated Bacterial Proteins. *Cell Host Microbe*. 2009;5(4):397-403. doi:10.1016/j.chom.2009.03.004
  26. Yan D, Quan H, Wang L, Liu F, Liu H, Chen J, Cao X, Ge B. Enteropathogenic *Escherichia coli* Tir recruits cellular SHP-2 through ITIM motifs to suppress host immune response. *Cell Signal*. 2013 Sep;25(9):1887-94. doi: 10.1016/j.cellsig.2013.05.020.
  27. Knutton S, Baldwin T, Williams PH, McNeish AS. Actin accumulation at sites of bacterial adhesion to tissue culture cells: Basis of a new diagnostic test for enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infect Immun*. 1989;57(4):1290-1298. doi:10.1128/iai.57.4.1290-1298.1989
  28. Luo Y, Frey EA, Pfuetzner RA, et al. Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. *Nature*. 2000;405(6790):1073-1077. doi:10.1038/35016618
  29. Hartland EL, Batchelor M, Delahay RM, et al. Binding of intimin from enteropathogenic *Escherichia coli* to Tir and to host cells. *Mol Microbiol*. 1999;32(1):151-158. doi:10.1046/j.1365-2958.1999.01338.x
  30. Mészáros B, Erdős G, Dosztányi Z. IUPred2A: Context-dependent prediction of protein



- disorder as a function of redox state and protein binding. *Nucleic Acids Res.* 2018;46(W1):W329-W337. doi:10.1093/nar/gky384
31. Blasche S, Arens S, Ceol A, et al. The EHEC-host interactome reveals novel targets for the translocated intimin receptor. *Sci Rep.* 2014;4:22-26. doi:10.1038/srep07531
  32. Blank U, Launay P, Benhamou M, Monteiro RC. Inhibitory ITAMs as novel regulators of immunity. *Immunol Rev.* 2009;232(1):59-71. doi:10.1111/j.1600-065X.2009.00832.x
  33. Hayashi T, Senda M, Suzuki N, et al. Differential Mechanisms for SHP2 Binding and Activation Are Exploited by Geographically Distinct *Helicobacter pylori* CagA Oncoproteins. *Cell Rep.* 2017;20(12):2876-2890. doi:10.1016/j.celrep.2017.08.080
  34. Hayward RD, Leong JM, Koronakis V, Campellone KG. Exploiting pathogenic *Escherichia coli* to model transmembrane receptor signalling. *Nat Rev Microbiol.* 2006;4(5):358-370. doi:10.1038/nrmicro1391
  35. Smith K, Humphreys D, Hume PJ, Koronakis V. Enteropathogenic *Escherichia coli* Recruits the Cellular Inositol Phosphatase SHIP2 to Regulate Actin-Pedestal Formation. *Cell Host Microbe.* 2010;7(1):13-24. doi:10.1016/j.chom.2009.12.004
  36. Phillips N, Hayward RD, Koronakis V. Phosphorylation of the enteropathogenic *E. coli* receptor by the Src-family kinase c-Fyn triggers actin pedestal formation. *Nat Cell Biol.* 2004;6(7):618-625. doi:10.1038/ncb1148
  37. Frese S, Schubert WD, Findeis AC, et al. The phosphotyrosine peptide binding specificity of Nck1 and Nck2 Src homology 2 domains. *J Biol Chem.* 2006;281(26):18236-18245. doi:10.1074/jbc.M512917200
  38. Lai Y, Rosenshine I, Leong JM, Frankel G. Intimate host attachment: Enteropathogenic and enterohaemorrhagic *Escherichia coli*. *Cell Microbiol.* 2013;15(11):1796-1808. doi:10.1111/cmi.12179
  39. De Groot JC, Schlüter K, Carius Y, et al. Structural basis for complex formation between human IRSp53 and the translocated intimin receptor tir of enterohemorrhagic *E. coli*. *Structure.* 2011;19(9):1294-1306. doi:10.1016/j.str.2011.06.015
  40. Yi C ryun, Goldberg MB. Putting Enterohemorrhagic *E. coli* on a Pedestal. *Cell Host Microbe.* 2009;5(3):215-217. doi:10.1016/j.chom.2009.03.002
  41. Beutler BA. TLRs and innate immunity. *Blood.* 2009;113(7):1399-1407. doi:10.1182/blood-2008-07-019307
  42. Takeuchi O, Akira S. Pattern Recognition Receptors and Inflammation. *Cell.* 2010;140(6):805-820. doi:10.1016/j.cell.2010.01.022
  43. Palm NW, Medzhitov R. Pattern recognition receptors and control of adaptive immunity. *Immunol Rev.* 2009;227(1):221-233. doi:10.1111/j.1600-065X.2008.00731.x
  44. van Hoof C, Goris J, Merlevede W. *Protein Phosphatases.*; 2007. doi:10.1002/9783527615032.ch12
  45. Ardito F, Giuliani M, Perrone D, Troiano G, Muzio L Lo. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). *Int J Mol Med.* 2017;40(2):271-280. doi:10.3892/ijmm.2017.3036

46. Neel BG, Gu H, Pao L. The 'Shp'ing news: SH2 domain-containing tyrosine phosphatases in cell signaling. *Trends Biochem Sci.* 2003;28(6):284-293. doi:10.1016/S0968-0004(03)00091-4
47. Cohen PTW. 1 Overview of protein serine / threonine phosphatases. 5(Table 1).
48. Massa PT, Ropka SL, Saha S, Fecenko KL, Beuler KL. Critical Role for Protein Tyrosine Phosphatase SHP-1 in Controlling Infection of Central Nervous System Glia and Demyelination by Theiler's Murine Encephalomyelitis Virus. *J Virol.* 2002;76(16):8335-8346. doi:10.1128/jvi.76.16.8335-8346.2002
49. Horvat A, Schwaiger FW, Hager G, et al. A novel role for protein tyrosine phosphatase SHP1 in controlling glial activation in the normal and injured nervous system. *J Neurosci.* 2001;21(3):865-874. doi:10.1523/jneurosci.21-03-00865.2001
50. Lorenz U. SHP-1 and SHP-2 in T cells: Two phosphatases functioning at many levels. *Immunol Rev.* 2009;228(1):342-359. doi:10.1111/j.1600-065X.2008.00760.x
51. Coggeshall KM, Nakamura K, Phee H. How do inhibitory phosphatases work? *Mol Immunol.* 2002;39(9):521-529. doi:10.1016/S0161-5890(02)00209-2
52. Yang J, Liu L, He D, et al. Crystal structure of human protein-tyrosine phosphatase SHP-1. *J Biol Chem.* 2003;278(8):6516-6520. doi:10.1074/jbc.M210430200
53. Carretero M, Palmieri G, Llano M, et al. Specific engagement of the CD94/NKG2-A killer inhibitory receptor by the HLA-E class Ib molecule induces SHP-1 phosphatase recruitment to tyrosine-phosphorylated NKG2-A: Evidence for receptor function in heterologous transfectants. *Eur J Immunol.* 1998;28(4):1280-1291. doi:10.1002/(SICI)1521-4141(199804)28:04<1280::AID-IMMU1280>3.0.CO;2-O
54. Campbell KS, Dessing M, Lopez-Botet M, Cella M, Colonna M. Tyrosine phosphorylation of a human killer inhibitory receptor recruits protein tyrosine phosphatase 1C. *J Exp Med.* 1996;184(1):93-100. doi:10.1084/jem.184.1.93
55. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol.* 2015;16(1):18-29. doi:10.1038/nrm3920
56. Delaforge, E., Cordeiro, T. N., Bernadó, P. & Sibille, N. Conformational characterization of intrinsically disordered proteins and its biological significance. in *Modern Magnetic Resonance* 381–399 (Springer International Publishing, 2018). doi:10.1007/978-3-319-28388-3\_52
57. Fleming PJ, Rose GD. Conformational Properties of Unfolded Proteins. *Protein Sci Encycl.* 2008:710-736. doi:10.1002/9783527610754.sf06
58. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol.* 2005;6(3):197-208. doi:10.1038/nrm1589
59. Habchi J, Tompa P, Longhi S, Uversky VN. Introducing protein intrinsic disorder. *Chem Rev.* 2014;114(13):6561-6588. doi:10.1021/cr400514h
60. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradović Z. Intrinsic disorder and protein function. *Biochemistry.* 2002;41(21):6573-6582. doi:10.1021/bi012159+
61. Iakoucheva LM, Radivojac P, Brown CJ, et al. The importance of intrinsic disorder for

- protein phosphorylation. *Nucleic Acids Res.* 2004;32(3):1037-1049. doi:10.1093/nar/gkh253
62. Wright PE, Dyson HJ. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J Mol Biol.* 1999;293(2):321-331. doi:10.1006/jmbi.1999.3110
  63. Uversky VN. The mysterious unfoldome: Structureless, underappreciated, yet vital part of any given proteome. *J Biomed Biotechnol.* 2010;2010. doi:10.1155/2010/568068
  64. Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, Dunker AK. Intrinsic disorder and functional proteomics. *Biophys J.* 2007;92(5):1439-1456. doi:10.1529/biophysj.106.094045
  65. Babu MM, van der Lee R, de Groot NS, Gsponer J. Intrinsically disordered proteins: Regulation and disease. *Curr Opin Struct Biol.* 2011;21(3):432-440. doi:10.1016/j.sbi.2011.03.011
  66. Tompa P. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett.* 2005;579(15):3346-3354. doi:10.1016/j.febslet.2005.03.072
  67. Felli IC, Pierattelli R, Tompa P. NMR to characterize flexible, disordered protein states: towards expanding the structure – function paradigm A CON 2D NMR experiment to highlight the importance of NMR to achieve atomic resolution information on intrinsically disordered proteins.
  68. Gołowicz D, Kasprzak P, Orekhov V, Kazimierczuk K. Fast time-resolved NMR with non-uniform sampling. *Prog Nucl Magn Reson Spectrosc.* 2020;116:40-55. doi:10.1016/j.pnmrs.2019.09.003
  69. Felli IC, Pierattelli R. Novel methods based on <sup>13</sup>C detection to study intrinsically disordered proteins. *J Magn Reson.* 2014;241(1):115-125. doi:10.1016/j.jmr.2013.10.020
  70. Kim S, Wu KP, Baum J. Fast hydrogen exchange affects <sup>15</sup>N relaxation measurements in intrinsically disordered proteins. *J Biomol NMR.* 2013;55(3):249-256. doi:10.1007/s10858-013-9706-1
  71. Wüthrich K. NMR studies of structure and function of biological macromolecules (Nobel Lecture). *Angew Chemie - Int Ed.* 2003;42(29):3340-3363. doi:10.1002/anie.200300595
  72. Purslow JA, Khatiwada B, Bayro MJ, Venditti V. NMR Methods for Structural Characterization of Protein-Protein Complexes. *Front Mol Biosci.* 2020;7. doi:10.3389/fmolb.2020.00009
  73. Redfield C. Resonance assignments in proteins. 1990:1-23.
  74. Sattler M, Schleucher J, Griesinger C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Magn Reson Spectrosc.* 1999;34(2):93-158. doi:10.1016/S0079-6565(98)00025-9
  75. Felli CI, Pierattelli R. Intrinsically Disordered Proteins Studied by NMR Spectroscopy. *Advances in Exp. Medicine and Biology* 870 Springer International Publishing, 2014)
  76. Kjaergaard M, Poulsen FM. Disordered proteins studied by chemical shifts. *Prog Nucl*

- Magn Reson Spectrosc.* 2012;60:42-51. doi:10.1016/j.pnmrs.2011.10.001
77. Shen Y, Lange O, Delaglio F, et al. Consistent blind protein structure generation from NMR chemical shift data. *Proc Natl Acad Sci U S A.* 2008;105(12):4685-4690. doi:10.1073/pnas.0800256105
  78. Rosenl w J, Isaksson L, Mayzel M, Lengqvist J, Orekhov VY. Tyrosine phosphorylation within the intrinsically disordered cytosolic domains of the B-cell receptor: An NMR-based structural analysis. *PLoS One.* 2014;9(4). doi:10.1371/journal.pone.0096199
  79. Wishart DS, Sykes BD. The <sup>13</sup>C Chemical-Shift Index: A simple method for the identification of protein secondary structure using <sup>13</sup>C chemical-shift data. *J Biomol NMR.* 1994;4(2):171-180. doi:10.1007/BF00175245
  80. Wishart DS, Sykes BD, Richards FM. The Chemical Shift Index: A Fast and Simple Method for the Assignment of Protein Secondary Structure Through NMR Spectroscopy. *Biochemistry.* 1992;31(6):1647-1651. doi:10.1021/bi00121a010
  81. Cavalli A, Salvatella X, Dobson CM, Vendruscolo M. Protein structure determination from NMR chemical shifts. *Proc Natl Acad Sci U S A.* 2007;104(23):9615-9620. doi:10.1073/pnas.0610313104
  82. Tamiola K, Mulder FAA. Using NMR chemical shifts to calculate the propensity for structural order and disorder in proteins. *Biochem Soc Trans.* 2012;40(5):1014-1020. doi:10.1042/BST20120171
  83. Theillet FX, Smet-Nocca C, Liokatis S, et al. Cell signaling, post-translational protein modifications and NMR spectroscopy. *J Biomol NMR.* 2012;54(3):217-236. doi:10.1007/s10858-012-9674-x
  84. Amata I, Maffei M, Igea A, et al. Multi-phosphorylation of the intrinsically disordered unique domain of c-src studied by in-cell and real-time NMR spectroscopy. *ChemBioChem.* 2013;14(14):1820-1827. doi:10.1002/cbic.201300139
  85. Du JT, Li YM, Wei W, et al. Low-barrier hydrogen bond between phosphate and the amide group in phosphopeptide. *J Am Chem Soc.* 2005;127(47):16350-16351. doi:10.1021/ja054568p
  86. Bienkiewicz EA, Lumb KJ. Random-coil chemical shifts of phosphorylated amino acids. *J Biomol NMR.* 1999;15(3):203-206. doi:10.1023/A:1008375029746
  87. Liokatis S, Dose A, Schwarzer D, Selenko P. Simultaneous detection of protein phosphorylation and acetylation by high-resolution NMR spectroscopy. *J Am Chem Soc.* 2010;132(42):14704-14705. doi:10.1021/ja106764y
  88. Selenko P, Frueh DP, Elsaesser SJ, Haas W, Gygi SP, Wagner G. In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. *Nat Struct Mol Biol.* 2008;15(3):321-329. doi:10.1038/nsmb.1395
  89. Theillet FX, Rose HM, Liokatis S, et al. Site-specific NMR mapping and time-resolved monitoring of serine and threonine phosphorylation in reconstituted kinase reactions and mammalian cell extracts. *Nat Protoc.* 2013;8(7):1416-1432. doi:10.1038/nprot.2013.083
  90. Theillet FX, Liokatis S, Jost JO, et al. Site-specific mapping and time-resolved monitoring

- of lysine methylation by high-resolution NMR spectroscopy. *J Am Chem Soc.* 2012;134(18):7616-7619. doi:10.1021/ja301895f
91. Zuiderweg ERP. Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry.* 2002;41(1):1-7. doi:10.1021/bi011870b
  92. Mollica L, Bessa LM, Hanouille X, Jensen MR, Blackledge M, Schneider R. Binding mechanisms of intrinsically disordered proteins: Theory, simulation, and experiment. *Front Mol Biosci.* 2016;3(SEP):1-18. doi:10.3389/fmolb.2016.00052
  93. Teilum K, Kunze MBA, Erlandsson S, Kragelund BB. (S)Pinning down protein interactions by NMR. *Protein Sci.* 2017;26(3):436-451. doi:10.1002/pro.3105
  94. Williamson MP. Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc.* 2013;73:1-16. doi:10.1016/j.pnmrs.2013.02.001
  95. Driscoll PC. Macromolecular Complexes. *Protein NMR Spectrosc Pract Tech Appl.* 2011:269-317. doi:10.1002/9781119972006.ch8
  96. Granot J. Determination of dissociation constants of 1:1 complexes from NMR data. Optimization of the experimental setup by statistical analysis of simulated experiments. *J Magn Reson.* 1983;55(2):216-224. doi:10.1016/0022-2364(83)90234-2
  97. Falconer RJ. Applications of isothermal titration calorimetry - the research and technical developments from 2011 to 2015. *J Mol Recognit.* 2016;29(10):504-515. doi:10.1002/jmr.2550
  98. LaVallie, E., DiBlasio, E., Kovacic, S. *et al.* A Thioredoxin Gene Fusion Expression System That Circumvents Inclusion Body Formation in the *E. coli* Cytoplasm. *Nat Biotechnol* **11**, 187–193 (1993).
  99. Smith DB, Johnson KS. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene.* 1988;67(1):31-40. doi:10.1016/0378-1119(88)90005-4
  100. Keller R. *The Computer Aided Resonance Assignment Tutorial.*; 2004.
  101. Tamiola K, Acar B, Mulder FAA. Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J Am Chem Soc.* 2010;132(51):18000-18003. doi:10.1021/ja105656t
  102. Livernois AM, Hnatchuk DJ, Findlater EE, Graether SP. Obtaining highly purified intrinsically disordered protein by boiling lysis and single step ion exchange. *Anal Biochem.* 2009;392(1):70-76. doi:10.1016/j.ab.2009.05.023
  103. Schmidt TGM, Skerra A. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc.* 2007;2(6):1528-1535. doi:10.1038/nprot.2007.209
  104. Stormo GD, Schneider TD, Gold LM. Characterization of translational initiation sites in *E. coli*. *Nucleic Acids Res.* 1982 May 11;10(9):2971-96. doi: 10.1093/nar/10.9.2971.
  105. Carter P, Kelley RF, Rodrigues ML, et al. High level *Escherichia coli* expression and production of a bivalent humanized antibody fragment. *Bio/Technology.* 1992;10(2):163. doi:10.1038/nbt0292-163

106. De Marco V, Stier G, Blandin S, De Marco A. The solubility and stability of recombinant proteins are increased by their fusion to NusA. *Biochem Biophys Res Commun.* 2004;322(3):766-771. doi:10.1016/j.bbrc.2004.07.189
107. Stevenson J, Krycer JR, Phan L, Brown AJ. A practical comparison of ligation-independent cloning techniques. *PLoS One.* 2013;8(12):8-14. doi:10.1371/journal.pone.0083888
108. Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res.* 1990;18(20):6069-6074. doi:10.1093/nar/18.20.6069
109. Berrow NS, Alderton D, Sainsbury S, et al. A versatile ligation-independent cloning method suitable for high-throughput expression screening applications. *Nucleic Acids Res.* 2007;35(6). doi:10.1093/nar/gkm047
110. Weeks SD, Drinker M, Loll PJ. Ligation independent cloning vectors for expression of SUMO fusions. *Protein Expr Purif.* 2007;53(1):40-50. doi:10.1016/j.pep.2006.12.006
111. Glick BR. Metabolic load and heterologous gene expression. *Biotechnol Adv.* 1995;13(2):247-261. doi:10.1016/0734-9750(95)00004-A
112. Donovan RS, Robinson CW, Click BR. Review: Optimizing inducer and culture conditions for expression of foreign proteins under the control of the lac promoter. *J Ind Microbiol.* 1996;16(3):145-154. doi:10.1007/BF01569997
113. Hochuli E, Bannwarth W, Dobeli H, Gentzi R, Stuber D. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Bio/Technology.* 1988;6(11):1321-1325. doi:10.1038/nbt1188-1321
114. Warawa J, Kenny B. Phosphoserine modification of the enteropathogenic Escherichia coli Tir molecule is required to trigger conformational changes in Tir and efficient pedestal elongation. *Mol Microbiol.* 2001;42(5):1269-1280. doi:10.1046/j.1365-2958.2001.02693.x
115. Xu C, Gagnon E, Call ME, et al. Regulation of T Cell Receptor Activation by Dynamic Membrane Binding of the CD3 $\epsilon$  Cytoplasmic Tyrosine-Based Motif. *Cell.* 2008;135(4):702-713. doi:10.1016/j.cell.2008.09.044
116. Cheng H, Schwell V, Curtis BR, Fazlieva R, Roder H, Campbell KS. Conformational Changes in the Cytoplasmic Region of KIR3DL1 upon Interaction with SHP-2. *Structure.* 2019;27(4):639-650.e2. doi:10.1016/j.str.2019.01.009
117. Salazar C, Höfer T. Multisite protein phosphorylation - From molecular mechanisms to kinetic models. *FEBS J.* 2009;276(12):3177-3198. doi:10.1111/j.1742-4658.2009.07027.x
118. Prévost C, Zhao H, Manzi J, et al. IRSp53 senses negative membrane curvature and phase separates along membrane tubules. *Nat Commun.* 2015;6. doi:10.1038/ncomms9529
119. Mylona A, Theillet FX, Foster C, et al. Opposing effects of Elk-1 multisite phosphorylation shape its response to ERK activation. *Science (80- ).* 2016;354(6309):233-237. doi:10.1126/science.aad1872
120. Bloch V, Yang Y, Margeat E, et al. The H-NS dimerization domain defines a new fold contributing to DNA recognition. *Nat Struct Biol.* 2003;10(3):212-218. doi:10.1038/nsb904
121. Fernando T, Royer C. Role of protein--protein interactions in the regulation of transcription

- by trp repressor investigated by fluorescence spectroscopy. *Biochemistry*. 1992;31(13):3429-3441. <http://www.ncbi.nlm.nih.gov/pubmed/1554725>.
122. De Biasio A, De Opakua AI, Mortuza GB, et al. Structure of p15PAF-PCNA complex and implications for clamp sliding during DNA replication and repair. *Nat Commun*. 2015;6. doi:10.1038/ncomms7439
123. Jensen MR, Communie G, Ribeiro EA, et al. Intrinsic disorder in measles virus nucleocapsids. *Proc Natl Acad Sci U S A*. 2011;108(24):9839-9844. doi:10.1073/pnas.1103270108
124. Schneider R, Maurin D, Communie G, et al. Visualizing the molecular recognition trajectory of an intrinsically disordered protein using multinuclear relaxation dispersion NMR. *J Am Chem Soc*. 2015;137(3):1220-1229. doi:10.1021/ja511066q



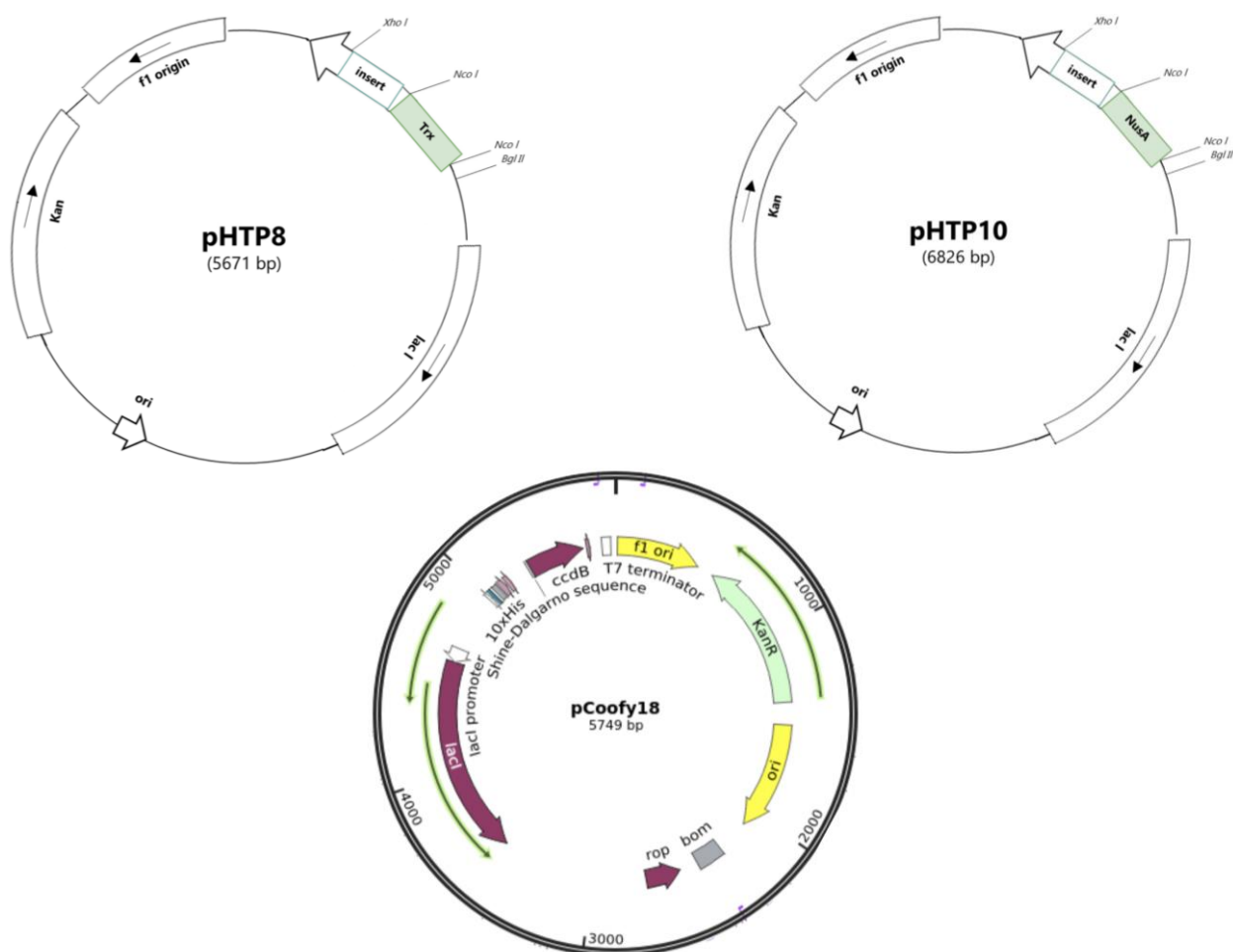


## **7. Anexos**



**Tabela 7.1- Informação das proteínas C-Tir, N-SH2, C-SH2 e NC-SH2 abordadas nesta dissertação.**

| Proteína | Gene  | Massa molecular (kDa) | Coefficiente de extinção molar ( $\epsilon$ ) ( $M^{-1}cm^{-1}$ ) |
|----------|---|-----------------------|---|
| C-Tir    | RRNQPAEQTTTTTHTVVQQQTGGNTPAQGGTDATR<br>AEDASLNRRDSQGSVASTHWSDDSSSEVVNPYAEVGGA<br>RNSLSAHQPEEHYDEVAADPGYSVIQNFSGSGPVTGR<br>LIGTPGQGIQSTYALLANSGLRLGMGGLTSGGESAVS<br>SVNAAPTGPVRFV  | 17,88                 | 11460   |
| N-SH2    | MVRWFHRDLSGLDAETLLKGRGVHGSFLARPSRKNQ<br>GDFLSVRVGDQVTHIRIQNSGDFYDLYGGEKFATLTE<br>LVEYYTQQQGVQLQDRDGTIIHLKYPLNCSDPTSWSH<br>PQFEK   | 13,84                 | 18450   |
| C-SH2    | SDPTSERWYHGHMSGGQAETLLQAKGEPWTFVLVRES<br>LSQPGDFVLSVLSDQPKAGPGSPLRVTHIKVMCEGGRY<br>TVGGLETFDLTDLVEHFKKTGIEEASGAFVYLRQPYYAT<br>RWSHPQFEK   | 14,44                 | 23950   |
| NC-SH2   | MVRWFHRDLSGLDAETLLKGRGVHGSFLARPSRKNQ<br>DFLSVRVGDQVTHIRIQNSGDFYDLYGGEKFATLTELV<br>EYYTQQQGVQLQDRDGTIIHLKYPLNCSDPTSERWYHGH<br>MSGGQAETLLQAKGEPWTFVLVRESLSQPGDFVLSVLSD<br>QPKAGPGSPLRVTHIKVMCEGGRYTVGGLETFDLTDL<br>VEHFKKTGIEEASGAFVYLRQPYYATRWS HPQFEK | 26,05                 | 36900   |



**Figura 7.1- Vetores de expressão nos quais foram clonados os genes das proteínas C-Tir, N-SH2, C-SH2 e NC-SH2**

**Tabela 7.2- Programas de amplificação dos genes N-SH2, C-SH2, NC-SH2 de SHP-1 para clonagem das diferentes construções nos vetores pHTP8 pHTP10 e pCoofy18, através de PCR.**

| Domínio de interesse | C-SH2  |       | N-SH2  |       |        |       | NC-SH2 |       |        |       |
|----------------------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|
| Vetor                | pHTP8  |       | pHTP8  |       | pHTP10 |       | pHTP8  |       | pHTP10 |       |
| Condições            | T (°C) | t (s) | T (°C) | t (s) | T (°C) | t (s) | T (°C) | t (s) | T (°C) | t (s) |
| Desnaturação inicial | 95     | 120   | 95     | 120   | 95     | 120   | 95     | 120   | 95     | 120   |
| Desnaturação         | 95     | 30    | 95     | 30    | 95     | 30    | 95     | 30    | 95     | 30    |
| <i>Annealing</i>     | 53,7   | 45    | 51,4   | 45    | 51,4   | 45    | 53,7   | 45    | 53,7   | 45    |
| Extensão             | 72     | 24    | 72     | 23    | 72     | 23    | 72     | 42    | 72     | 42    |
| Entensão final       | 72     | 300   | 72     | 300   | 72     | 300   | 72     | 300   | 72     | 300   |

| Domínio de interesse | N-SH2            |       |               |       | NC-SH2           |       |               |       |
|----------------------|------------------|-------|---------------|-------|------------------|-------|---------------|-------|
| Vetor                | pHTP8 (c/linker) |       | pHTP8 (c/TEV) |       | pHTP8 (c/linker) |       | pHTP8 (c/TEV) |       |
| Condições            | T (°C)           | t (s) | T (°C)        | t (s) | T (°C)           | t (s) | T (°C)        | t (s) |
| Desnaturação inicial | 95               | 120   | 95            | 120   | 95               | 120   | 95            | 120   |
| Desnaturação         | 95               | 30    | 95            | 30    | 95               | 30    | 95            | 30    |
| <i>Annealing</i>     | 49,8             | 45    | 49,8          | 45    | 52,1             | 45    | 52,1          | 45    |
| Extensão             | 72               | 24    | 72            | 24    | 72               | 45    | 72            | 45    |
| Entensão final       | 72               | 300   | 72            | 300   | 72               | 300   | 72            | 300   |

| Domínio de interesse | N-SH2    |       | NC-SH2   |       |
|----------------------|----------|-------|----------|-------|
| Vetor                | pCoofy18 |       | pCoofy18 |       |
| Condições            | T (°C)   | t (s) | T (°C)   | t (s) |
| Desnaturação inicial | 95       | 120   | 95       | 120   |
| Desnaturação         | 95       | 30    | 95       | 30    |
| <i>Annealing</i>     | 69       | 45    | 70       | 45    |
| Extensão             | 72       | 23    | 72       | 42    |
| Entensão final       | 72       | 300   | 72       | 300   |

**Tabela 7.3- Componentes da amplificação dos genes N-SH2, C-SH2, NC-SH2 de SHP-1 para clonagem das diferentes construções nos vetores pHTP8, pHTP10 e pCoofy18, através de PCR.**

| Vetor                      | pHTP8 e pHTP10 | Vetor                  | pCoofy18  |
|----------------------------|----------------|------------------------|-----------|
| NZY Proof Green M. mix 25x | 25 µL          | Fusion High Fidelity   | 0,5 µL    |
| Primer Forward (10 µM)     | 1,25 µL        | Primer Forward (10 µM) | 2,5 µL    |
| Primer Reverse (10 µM)     | 1,25 µL        | Primer Reverse (10 µM) | 2,5 µL    |
| DNA molde (0,5 ng)         | 1,25 µL        | DNA molde (0,5 ng)     | 1 µL      |
| H <sub>2</sub> O           | até 50 µL      | HF <i>buffer</i>       | 10 µL     |
|                            |                | H <sub>2</sub> O       | até 50 µL |

**Tabela 7.4- Componentes e programa da reação de clonagem dos genes N-SH2, C-SH2, NC-SH2 de SHP-1 para clonagem das diferentes construções nos vetores pHTP8 pHTP10, através de PCR.**

| Vetor               |  | pHTP8 e pHTP10                     |  | T (°C) | t (min) |
|---------------------|--|------------------------------------|--|--------|---------|
| NZY easy enzyme mix |  | 0,5 µL                             |  | 37     | 60      |
| DNA                 |  | (pb fragmento amplificado x 0,083) |  | 80     | 10      |
| Vetor linearizado   |  | 1 µL                               |  | 30     | 10      |
| Mistura reacional   |  | 1 µL                               |  |        |         |
| H <sub>2</sub> O    |  | até 10 µL                          |  |        |         |

**Tabela 7.5- Componentes e programa da reação de clonagem dos genes N-SH2 e NC-SH2 de SHP-1 para clonagem no vetor pCoofy18, através de PCR.**

| Vetor             |  | pCoofy18     |  | T (°C) | t (min) |
|-------------------|--|--------------|--|--------|---------|
| Enzima recA       |  | 0,5 µL       |  | 37     | 30      |
| DNA               |  | 1 µL (7 ng)  |  |        |         |
| Vetor linearizado |  | 1 µL (30 ng) |  |        |         |
| Tampão recA       |  | 1 µL         |  |        |         |
| H <sub>2</sub> O  |  | até 10 µL    |  |        |         |

**Tabela 7.6- Programas do *screening* das colônias recombinantes dos genes N-SH2, C-SH2, NC-SH2 de SHP-1 após clonagem das diferentes construções nos vetores pHTP8 pHTP10.**

| Domínios de interesse | C-SH2  |       | N-SH2  |       |        |       | NC-SH2 |       |        |       |
|-----------------------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|
|                       | pHTP8  |       | pHTP8  |       | pHTP10 |       | pHTP8  |       | pHTP10 |       |
| Condições             | T (°C) | t (s) | T (°C) | t (s) | T (°C) | t (s) | T (°C) | t (s) | T (°C) | t (s) |
| Desnaturação inicial  | 95     | 120   | 95     | 120   | 95     | 120   | 95     | 120   | 95     | 120   |
| Desnaturação          | 95     | 50    | 95     | 50    | 95     | 50    | 95     | 50    | 95     | 50    |
| <i>Annealing</i>      | 53,7   | 60    | 51,4   | 60    | 51,4   | 60    | 53,7   | 60    | 53,7   | 60    |
| Extensão              | 72     | 24    | 72     | 23    | 72     | 23    | 72     | 42    | 72     | 42    |
| Entensão final        | 72     | 300   | 72     | 300   | 72     | 300   | 72     | 300   | 72     | 300   |

| Domínios de interesse | N-SH2           |       |                  |       | NC-SH2          |       |                  |       |
|-----------------------|-----------------|-------|------------------|-------|-----------------|-------|------------------|-------|
|                       | pHTP8<br>(c/L8) |       | pHTP8<br>(c/TEV) |       | pHTP8<br>(c/L8) |       | pHTP8<br>(c/TEV) |       |
| Condições             | T (°C)          | t (s) | T (°C)           | t (s) | T (°C)          | t (s) | T (°C)           | t (s) |
| Desnaturação inicial  | 95              | 120   | 95               | 120   | 95              | 120   | 95               | 120   |
| Desnaturação          | 95              | 50    | 95               | 50    | 95              | 50    | 95               | 50    |
| <i>Annealing</i>      | 49,8            | 60    | 49,8             | 60    | 49,8            | 60    | 49,8             | 60    |
| Extensão              | 72              | 24    | 72               | 24    | 72              | 45    | 72               | 45    |
| Entensão final        | 72              | 300   | 72               | 300   | 72              | 300   | 72               | 300   |

**Tabela 7.7- Componentes do *screening* das colônias recombinantes dos genes alvo N-SH2, C-SH2, NC-SH2 de SHP-1 após clonagem das diferentes construções nos vetores pHTP8 pHTP10.**

| Vetor                    | pHTP8 e pHTP10 |
|--------------------------|----------------|
| NZYtech Taq Green M. mix | 12,5 µL        |
| Primer Forward (0,25 µM) | 1,25 µL        |
| Primer Reverse (0,25 µM) | 1,25 µL        |
| DNA                      | 1 µL           |
| H <sub>2</sub> O         | até 25 µL      |

**Tabela 7.8- Primers utilizados na amplificação dos genes alvo N-SH2, C-SH2, NC-SH2. A negrito está representada a *non-priming overhang*.**

| Gene   | Construção idealizada  | Primers        | Sequência de nucleótidos dos <i>primers</i>   |
|--------|--|----------------|---|
| N-SH2  | Trx-His <sub>6</sub> -N-SH2 e NusA-His <sub>6</sub> -N-SH2   | <i>Forward</i> | 5'- <b>TCAGCAAGGGCTGAGG</b> CTGGAAGTTCT GTTCCAGGGG-3'   |
|        |  | <i>Reverse</i> | 5'- <b>TCAGCGGAAGCTGAGG</b> CTACTTCTCGAA CTGCGGATGGCTCCAAGTCTGGGATCGGA GCA-3'                           |
|        | Trx-His <sub>6</sub> -L8-N-SH2                               | <i>Forward</i> | 5'- <b>TCAGCAAGGGCTGAGG</b> CTCGAGGTACT ATTTCAAGGACCAGGTGGAGGCAGTGGCGG TGGCTCTATGGTGAGGTGGTTTCAC-3'     |
|        |  | <i>Reverse</i> | 5'- <b>TCAGCGGAAGCTGAGG</b> CTACTTCTCGAA CTGCGGATGGCTCCAAGTCTGGGATCGGA GCA-3'                           |
|        | Trx-His <sub>6</sub> -TEV-L8-N-SH2                           | <i>Forward</i> | 5'- <b>TCAGCAAGGGCTGAGG</b> GAGAACCTATAC TTCCAAGGCCGGCATAACCTCTCCGGAAGC CCCAAGGATGGTGAGGTGGTTTCACCGA-3' |
|        |  | <i>Reverse</i> | 5'- <b>TCAGCGGAAGCTGAGG</b> CTACTTCTCGAA CTGCGGATGGCTCCAAGTCTGGGATCGGA GCA-3'                           |
| C-SH2  | Trx-His <sub>6</sub> -C-SH2                                  | <i>Forward</i> | 5'- <b>TCAGCAAGGGCTGAGG</b> CTCGAGGTACTC TTTCAAGGACCAGCTGGGATCCGGGCGAGCT TCCGATCCCACTAGTGAGAGG-3'       |
|        |  | <i>Reverse</i> | 5'- <b>TCAGCGGAAGCTGAGG</b> CTACTTCTCGAA CTGCGGATGGCTCCACCTCGTGGCATAGTA CGGCTG-3'                       |
| NC-SH2 | Trx-His <sub>6</sub> -NC-SH2 e NusA-His <sub>6</sub> -NC-SH2 | <i>Forward</i> | 5'- <b>TCAGCAAGGGCTGAGG</b> CTGGAAGTTCTGT TCCA GGGG-3'  |
|        |  | <i>Reverse</i> | 5'- <b>TCAGCGGAAGCTGAGG</b> CTACTTCTCGAA CTGCGGATGGCTCCACCTCGTGGCATAGTA CGGCTG-3'                       |
|        | Trx-His <sub>6</sub> -L8-NC-SH2                              | <i>Forward</i> | 5'- <b>TCAGCAAGGGCTGAGG</b> CTCGAGGTACTAT TTCAAGGACCAGGTGGAGGCAGTGGCGGTG GCTCTATGGTGAGGTGGTTTCAC-3'     |
|        |  | <i>Reverse</i> | 5'- <b>TCAGCGGAAGCTGAGG</b> CTACTTCTCGAA CTGCGGATGGCTCCACCTCGTGGCATAGTA CGGCTG-3'                       |
|        | Trx-His <sub>6</sub> -TEV-L8-NC-SH2                          | <i>Forward</i> | 5'- <b>TCAGCAAGGGCTGAGG</b> GAGAACCTATACT TCCAAGGCCAGGCATAACCTCTCCGGAAG CCCAAGGATGGTGAGGTGGTTTCACCGA-3' |
|        |  | <i>Reverse</i> | 5'- <b>TCAGCGGAAGCTGAGG</b> CTACTTCTCGAA CTGCGGATGGCTCCACCTCGTGGCATAGTA CGGCTG-3'                       |

|        |                           |         |  |
|--------|---------------------------|---------|--|
| N-SH2  | His <sub>10</sub> -N-SH2  | Forward | 5'- <b>AAGTTCTGTTCCAGGGGCCCC</b> ATGGTGAGGTG<br>GTTTCACCGA-3'  |
|        |                           | Reverse | 5'- <b>CCCCAGAACATCAGGTTAATGGCG</b> CTACTTCT<br>CGAACTGCGGATGGCTCCAGTCTGAGGAGTTCA<br>GCGGGTACTTGAGGTG-3' |
| NC-SH2 | His <sub>10</sub> -NC-SH2 | Forward | 5'- <b>AAGTTCTGTTCCAGGGGCCCC</b> ATGGTGAGGTG<br>GTTTCACCGA-3'  |
|        |                           | Reverse | 5'- <b>CCCCAGAACATCAGGTTAATGGCG</b> CTACTTCT<br>CGAACTGCGGATGGCTCCACCTCGTGGCATAGT<br>ACGGCTG-3'          |

**Tabela 7.9- Composição do gel SDS-PAGE (12%).**

| Reagentes                                  | Gel de separação | Gel de empilhamento |
|--|------------------|---------------------|
| Água Milli-Q                               | 3,3 mL           | 3,4 mL              |
| Bis-acrilamida (30%)                       | 4,0 mL           | 0,85 mL             |
| 1.5 M Tris-HCl pH=6,8 / 1M Tris-HCl pH=8,0 | 2,5 mL           | 625 mL              |
| 10% (m/v) SDS                              | 100 µL           | 50 µL               |
| 10% (v/v) APS                              | 100 µL           | 50 µL               |
| TEMED                                      | 10 µL            | 5 µL                |

**Tabela 7.10- Composição do tampão de corrida Tris-Glicina para SDS-PAGE pH=8,3 (1L).**

| Reagentes         |         |
|-------------------|---------|
| Tris-Base (25 mM) | 3,03 g  |
| Glicina (192 mM)  | 14,42 g |
| 0,1% (m/v) SDS    | 10 mL   |
| Água Milli-Q      | até 1 L |

**Tabela 7.11- Composição do gel de agarose 1% (m/v).**

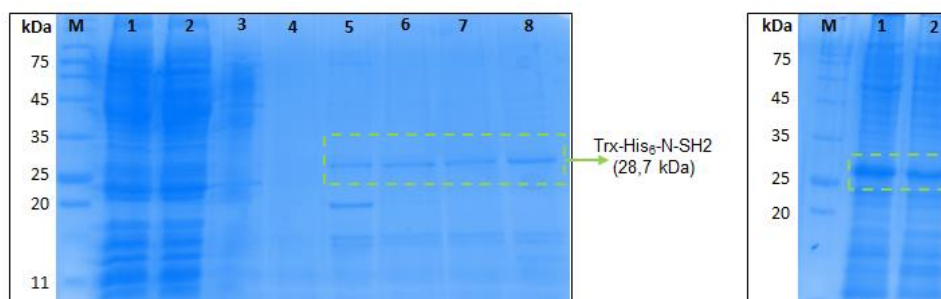
| Reagentes   |       |
|---|-------|
| Agarose (1%)  | 0,6 g |
| Tampão TAE (40 mM Tris-HCl pH=7,6;<br>20 mM ácido acético; 1 mM EDTA) | 60 mL |

**Tabela 7.12- Meio de cultura mínimo M9 para expressão de proteínas recombinantes isotopicamente marcadas.**

| Meio de cultura mínimo M9 (1 L)       |        | Trace elements 1 L (100x) pH=7,5     |        |
|---------------------------------------|--------|--------------------------------------|--------|
| Meio M9 (10x)                         | 100 mL | EDTA                                 | 5 g    |
| Trace elements (100x)                 | 10 mL  | FeCl <sub>3</sub> .6H <sub>2</sub> O | 0,83 g |
| Tiamina (1 mg/mL)                     | 1 mL   | ZnCl <sub>2</sub>                    | 84 mg  |
| Biotina (1 mg/mL)                     | 1 mL   | CuCl <sub>2</sub> .2H <sub>2</sub> O | 13 mg  |
| MgSO <sub>4</sub> (1M)                | 1 mL   | CoCl <sub>2</sub> .2H <sub>2</sub> O | 10 mg  |
| CaCl <sub>2</sub> (1M)                | 0,3 mL | H <sub>3</sub> BO <sub>3</sub>       | 10 mg  |
| 10% (m/v) Glucose                     | 20 mL  | MnCl <sub>4</sub> .H <sub>2</sub> O  | 1,6 mg |
| <sup>13</sup> C <sub>6</sub> -glucose | 2 g    |                                      |        |

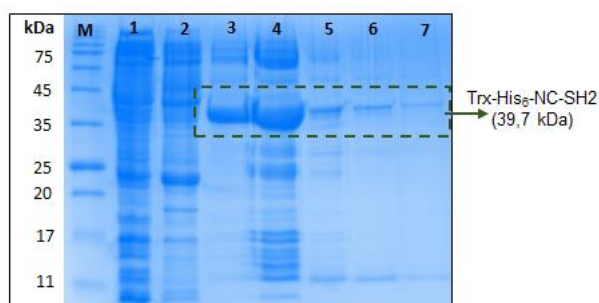
  

| Meio M9 1 L (10x) pH=7,2         |        |
|----------------------------------|--------|
| Na <sub>2</sub> HPO <sub>4</sub> | 75,2 g |
| KH <sub>2</sub> PO <sub>4</sub>  | 30 g   |
| NaCl                             | 5 g    |
| <sup>15</sup> NH <sub>4</sub> Cl | 5 g    |

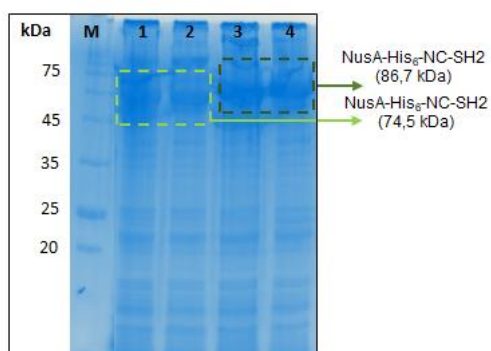


**Figura 7.2- SDS-PAGE (12%) da tentativa de purificação da construção Trx-His<sub>6</sub>-N-SH<sub>2</sub>. (A)- SDS-PAGE (12%) com as frações provenientes da cromatografia de afinidade realizada utilizando a coluna Strep-TactinXT™. M- Marcador de pesos moleculares NZYColour Protein Marker II; 1 e 2 - Fração *Flow-through*; 3- Fração de lavagem; 4 a 8- Frações de 5 mL de eluição; (B)- SDS-PAGE (12%) com as frações obtidas antes (poço 1) e depois (poço 2) da digestão com a protease HRV-3C.**



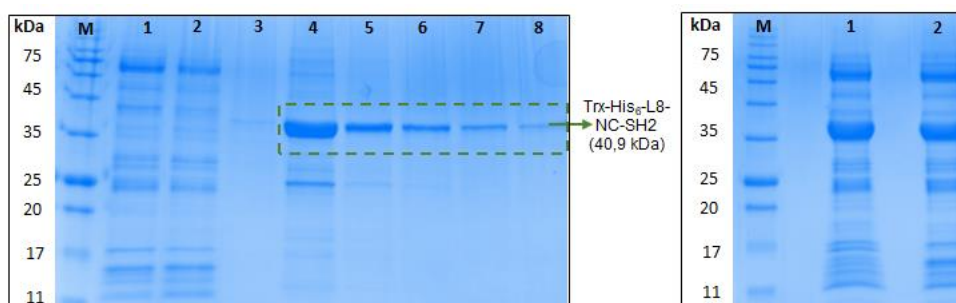


**Figura 7.3- SDS-PAGE (12%) da tentativa de purificação da construção Trx-His6-NC-SH2 com as frações provenientes da IMAC realizada utilizando a coluna HisTrap™.** M- Marcador de pesos moleculares NZYColour Protein Marker II; 1 Fração *Flow-through*; 2-Fração de lavagem; 3 a 7- Frações de 5 mL de eluição.

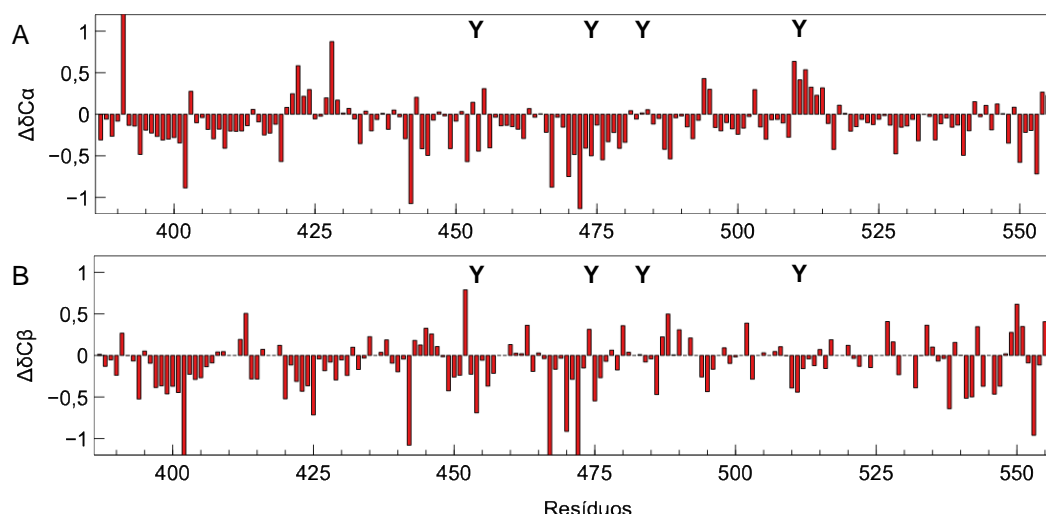


**Figura 7.4- SDS-PAGE (12%) da tentativa de purificação das construções NusA-His6-N-SH2 e NusA His6-NC-SH2.** M- Marcador de pesos moleculares NZYColour Protein Marker II; 1- Fração da construção NusA-His6-N-SH2 obtida antes da digestão com a protease HRV-3C; 2- Fração da construção NusA-His6-N-SH2 obtida depois da digestão com a protease HRV-3C; 3- Fração da construção NusA-His6-NC-SH2 obtida antes da digestão com a protease HRV-3C; 4- Fração da construção NusA-His6-NC-SH2 obtida depois da digestão com a protease HRV-3C.

O mesmo resultado foi obtido para as construções **Trx-His6-TEV-N-SH2 e Trx-His6-TEV-NC-SH2**, onde a digestão foi tentada com a protease TEV.



**Figura 7.5- SDS-PAGE (12%) da tentativa de purificação da construção Trx-His6-L8-NC-SH2. (A)-** SDS-PAGE (12%) com as frações provenientes da cromatografia de afinidade realizada utilizando a coluna Strep-TactinXT™. M- Marcador de pesos moleculares NZYColour Protein Marker II; 1 Fração *Flow-through*; 2-Fração de lavagem; 3 a 8- Frações de 5 mL de eluição; **(B)-** SDS-PAGE (12%) com as frações obtidas antes (poço 1) e depois (poço 2) da digestão com a protease HRV-3C.



**Figura 7.6- Desvios químicos secundários ( $\Delta\delta^{C\alpha}$  e  $\Delta\delta^{C\beta}$ ) calculados para C-Tir no estado fosforilado.** Os  $\Delta\delta^{C\alpha}$  (A) e  $\Delta\delta^{C\beta}$  (B) foram calculados através da diferença entre os desvios químicos Cα e Cβ experimentais e os desvios químicos referentes aos mesmos átomos corrigidos para IDPs, respetivamente. A posição das tirosinas fosforiladas na sequência de C-Tir encontram-se representadas pela letra Y.

### Cálculo do tempo de correlação rotacional ( $\tau_c$ )

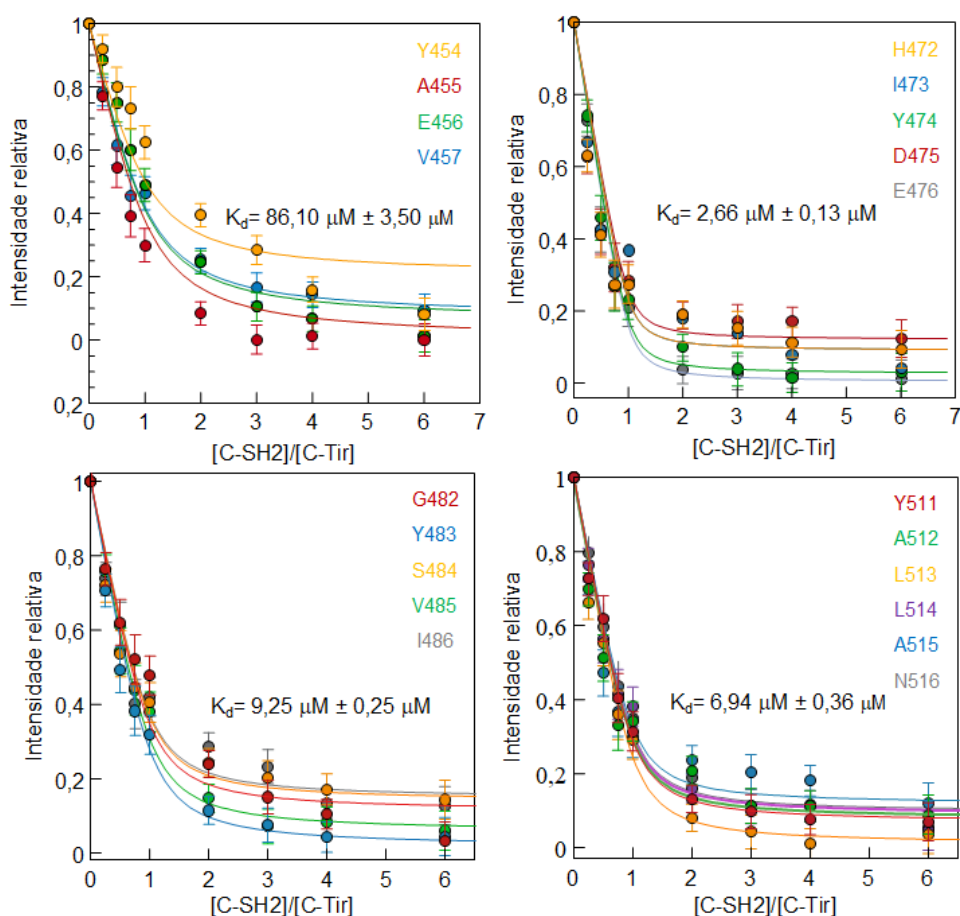
Assumindo uma forma esférica, o cálculo do tempo de correlação rotacional ( $\tau_c$ ) pode ser conseguido a partir da massa molecular (MW) do mesmo, através da equação 7.1.

$$\tau_c = \frac{MW\eta(V_p + \delta_w V_w)}{RT} \quad [7.1]$$

Onde  $V_p$  e  $V_w$  representam os volumes específicos da proteína e da água (estimados para 0,72 e 1, respetivamente).  $\delta_w$  representa o fator de hidratação da proteína (estimado para 0,35) e onde  $\eta$  representa a viscosidade do tampão à temperatura  $t$ , que pode ser obtida usando a equação 7.2:

$$1,7753 - 0,0565t + 1,0751 \times 10^{-3}t^2 - 9,2222 \times 10^{-6}t^3 \quad [7.2]$$

Uma vez calculado o  $\tau_c$  do domínio tendo em conta as condições da experiência, podemos através da mesma equação 7.1 calcular a massa molecular. Dessa forma, concluímos que o domínio C-SH2 às condições a que foi realizada a experiência comporta-se como uma entidade de 22,2 kDa (e não de 14,4 kDa, como se comportaria se a mesma fosse realizada a 25 °C).



**Figura 7.7- Curvas de titulação dos resíduos vizinhos e compreendidos nas quatro regiões ITIM presentes na sequência de C-Tir. (A)-** Curvas de titulação do ITIM que contém a Y454. **(B)-** Curvas de titulação do ITIM que contém a Y474. **(C)-** Curvas de titulação do ITIM que contém a Y483. **(D)-** Curvas de titulação do ITIM que contém a Y511. Os  $K_d$  indicam as forças de interação de cada região ITIM para o domínio C-SH2 de SHP-1. As barras de erros indicam o desvio padrão do  $F_{escala}$  calculado para cada espectro adquirido.